UNIVERSITÄT BERN

Graduate School for Cellular and Biomedical Sciences

University of Bern

Forward vs. reverse genetics: a bovine perspective based on visible and hidden phenotypes of inherited disorders

PhD Thesis submitted by

Irene Monika Häfliger

from **Fischbach LU**

for the degree of

PhD in Computational Biology

Supervisor

Prof. Dr. Cord Drögemüller

Institute of Genetics

Vetsuisse Faculty of the University of Bern

Co-advisor

Dr. Rémy Bruggmann

Interfaculty Bioinformatics Unit

Faculty of Science of the University of Bern

Accepted by the Faculty of Medicine, the Faculty of Science and the Vetsuisse Faculty of the University of Bern at the request of the Graduate School for Cellular and Biomedical Sciences

Bern, Dean of the Faculty of Medicine

Bern, Dean of the Faculty of Science

Bern, Dean of the Vetsuisse Faculty Bern

I Abstract

In modern cattle production, we have seen a negative trend for decades in reproduction while productivity and performance have improved. Although considered genetically complex, part of these fecundity, fertility, and rearing success issues are caused by Mendelian monogenic disorders. Traditionally, such disorders are investigated opportunistically based on their sporadic occurrence and through subsequent targeted analysis of affected individuals. This approach is called the **forward genetic approach (FGA)**. Modern genomic technologies, such as single nucleotide polymorphism (SNP) array genotyping and whole-genome sequencing (WGS), allow for straightforward locus mapping and the identification of candidate causal variants in affected individuals or families. Nevertheless, a major drawback is the arbitrary sampling and availability of well-phenotyped individuals for research, especially for mostly invisible defects affecting fecundity, early embryonic death, and abortions. Therefore, the **reverse genetic approach (RGA)** is applied to screen for underlying recessive lethal or sub-lethal variants. This approach requires the availability of massive population-wide genomic data. By applying a haplotype screen for a significant deviation of the Hardy-Weinberg equilibrium, genomic regions potentially harboring candidate causal variants are identified. The subsequent generation of WGS data of haplotype carriers allows for the mining for pathogenic variants potentially causing a reduction in homozygosity.

In the first part of my thesis, I present 18 successful, 1 inconclusive example, and 1 example addressing co-dominant effects of a known disorder. These FGA analyzes include heritable skin (n=7), bone (n=7), neuromuscular (n=1), eye (n=2), as well as syndromic disorders (n=3) in various European cattle breeds. Missense and frameshift variants in the *IL17RA*, *DSP,* and *FA2H* genes were described in three recessive genodermatoses: immunodeficiency with psoriasis-like skin alterations, syndromic ichthyosis, and ichthyosis congenita, respectively. Hypohidrotic ectodermal dysplasia was described as X-linked disorder that is associated with a gross deletion in the *EDA* gene. In dominant genodermatoses, a missense variant in *COL5A2* was shown to lead to classical Ehlers-Danlos syndrome, an in-frame deletion in *KRT5* was shown to cause epidermolysis bullosa simplex, and results of a study using an individual case of juvenile angiomatosis remained inconclusive. A recessive disorder described as hemifacial macrosomia was associated with a missense variant in *LAMB1*. Chondrodysplasia in a single family was shown to be caused by a *de novo* mutation in the bull leading to a stop-loss of the gene *FGFR3*. *De novo* mutations (missense and large deletions) in the *COL2A1* and *COL1A1* genes were associated with achondrogenesis type II (bulldog calf syndrome), and osteogenesis imperfecta type II, respectively. Another mutation that we

found to affect bone morphology was a trisomy in chromosome 29 leading to proportional dwarfism with facial dysplasia. Congenital neuromuscular channelopathy was for the first time associated with a missense variant in *KCNG1*. Furthermore, a *de novo* missense variant in *ADAMTSL4* and a recessive missense variant in *CNGB3* were shown to cause congenital cataract and achromatopsia, respectively. Additionally, cases of pulmonary hypoplasia and anasarca syndrome were analyzed and shown to be caused by trisomy 20 in two unrelated calves and a recessively inherited missense variant in *ADAMTS3*. Moreover, the fatal syndromic disorder skeletal-cardo-enteric dysplasia was described to be caused by a *de novo* missense variant in *MAP2K2*. Finally, I investigated the effects on blood cholesterol and triglyceride levels of heterozygous carriers of the previously described *APOB*-related cholesterol deficiency.

In the second part of my thesis, I present the outcome of the RGA in four main Swiss populations, that was validated with the SWISScow custom array. In the Brown Swiss dairy population, 72 haplotype regions showed significant depletions in homozygosity. Four of these haplotypes (BH6, BH14, BH24, and BH34) were associated with missense and nonsense variants in different genes (*MARS2*, *MRPL55*, *CPT1C*, and *ACSL5*, respectively). In the Original Braunvieh population, eight haplotype regions were identified. Candidate causal variants included a missense variant in *TUBGCP5* gene associated with haplotype OH2, and a splice site frameshift variant in *LIG3* gene associated with haplotype OH4. In the Holstein population, 24 haplotype regions were identified with a significant reduction of homozygosity. Subsequently, four novel candidate variants were proposed: a nonsense variant in *KIR2DS1* for haplotype HH13, in-frame deletion in the genes *NOTCH3* for HH21 haplotype, and *RIOX1* for HH25 haplotype, and finally, a missense variant in *PCDH15* for HH35 haplotype. In the Simmental population, eleven haplotype regions were detected. The haplotype SH5 was associated with a frameshift variant in *DIS3* gene and the haplotypes SH8 and SH9 with missense variants in the *CYP2B6* and *NUBPL* genes, respectively. For the breeds Brown Swiss, Original Braunvieh, and Holstein, association studies were carried out including traits describing fertility, birth, growth, and survival. Thereby most of the described mentioned haplotypes show additive effects.

Regardless of the approach, all the described candidate causal variants can be used as a tool of precision diagnostics and represent a step forward towards personalized medicine in cattle. Furthermore, these variants can be easily genotyped and allow for targeted breeding to reduce the number of risk matings, which would lead to a reduction of affected animals and significant improvement in animal health and welfare.

II Tables of contents, figures and tables

i Table of contents

ii List of figures

iii List of tables

iv List of additional in the appendix

1 Introduction

Cattle (*Bos taurus* and *Bos indicus*) were domesticated roughly 10,000 years ago and played a vital role in the process of nomadic humans becoming settlers [1]. While providing products such as milk, meat, and leather, in comparison to other livestock, cattle also provided draft power to plow fields and thereby facilitating pastoralism. Historical findings show the first evidence of pastoralism in the Swiss mountains about 4,200 years ago [2]. Then and now, especially for mountainous countries, ruminants are very important, as they can effectively digest rough plants without relying on food used for human diets. In 2015, the global cattle population was estimated to encompass 1.43 billion animals [3].

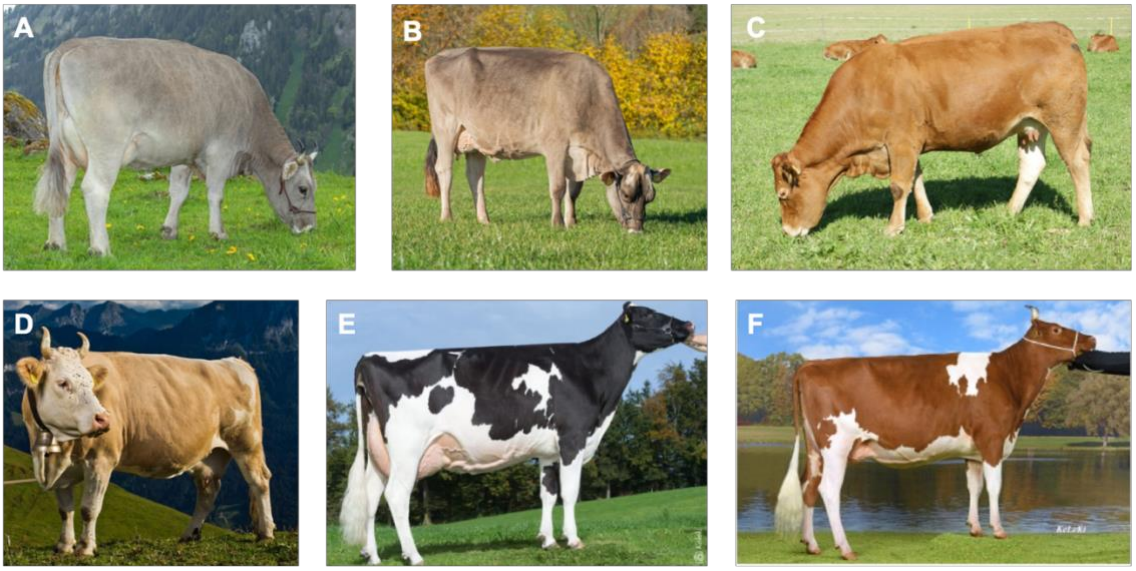


Figure 1: Typical cows representing the most important cattle populations in Switzerland (A) Original Braunvieh (dual-purpose) [4], (B) Brown Swiss (dairy) [4], (C) Limousin (beef) [5], (D) Simmental (dual-purpose) [6], (E) Holstein (dairy) [7] and (F) Swiss Fleckvieh (dairy) [8].

After a continuous decline over the last decade, the current Swiss cattle population includes 1,500,826 registered living animals, with the most prominent populations being the international breeds Holstein (HO) and Brown Swiss (BS), followed by crossbred animals [9]. Figure 1 illustrates the most important dairy, dual-purpose, and beef breeds in Switzerland, and Table 1 shows their population sizes. Regarding the two Swiss brown cattle populations, Original Braunvieh (OB) is the native Swiss brown breed and does not have any recent influence from other breeds, contrarily, BS is based on OB with introgression of international BS cattle [10]. Population analyzes showed that OB and BS are indeed genetically diverse and form two distinct populations [11, 12]. Also, the local Swiss Simmental (SI) population is a closed population native to Switzerland, more

precisely to the valley of Simmental in the canton of Bern. Interestingly, the breed Swiss Fleckvieh (SF), which arose from the crossbreeding of Holstein and Simmental, is clearly distinguishable from SI, but not from HO [11, 12].

Table 1: Population size of the main Swiss cattle breeds. Note the indication of the number of genotyped animals, representing the population included in the genomic breeding program.

breed	purpose	Swiss population[a]	animals in herdbook[b]	genotyped animals[c]	genotyped of Swiss population (%)
Holstein	dairy	401,616[d]	243,784[e]	58,564	14.58
Brown Swiss	dairy	263,283	147,877	55,382	21.04
Original Braunvieh	beef and dairy	45,161	13,149	11,574	25.63
Simmental	beef and dairy	93,679	23,102	10,539	11.25
Swiss Fleckvieh	dairy	135,958	64,651	*NA*	*NA*
Limousin	beef	92,247	2,805	5,366	5.82

[a]status as of 30th September 2021 [9]
[b]status as of 30th November 2020 [13–16]
[c]status as of 31st January 2021
[d]includes Holstein and Red Holstein
[e]includes animals from swissherdbook and Holstein Switzerland [13, 14]

1.1 The modern cattle breeding program

Traditional breeding programs had a major focus on production traits, such as milk production, milk composition, and slaughter weight. Modern breeding programs record more phenotypes population-wide and include both health-related traits as well as production traits. Thereby traits regarding fertility, birth, growth, and survival that indirectly describe rearing success are being included. For these trait groups, a list of specific traits that are part of the Swiss breeding programs can be found in Table A1, which are e.g., non-return rate as fertility trait, birth weight as birth trait, etc. With the application of quantitative genetic approaches to predict breeding values in the 1980ies, animals within a population can be compared on a statistical level by their performance [17]. The systematic use of estimated breeding values led to an incredible production increase over just a few decades. In order to select efficiently, the traits need to be heritable [17]. Only 10 years ago the implementation of genomic selection (GS) in breeding programs was proposed [18]. This method builds on the availability of single nucleotide polymorphism (SNP) array genotyping [18, 19]. From 2008, bovine SNP arrays were industrially available. Nevertheless, this brilliant idea was ahead of its time, as the introduction of GS relies on a comprehensive database of genotypes representing the base population [18]. In theory, GS is particularly interesting for multifactorial traits

with low heritability, because it accounts directly for Mendelian sampling [18, 19]. Regarding the application of GS, it is proven to improve the reliability of breeding values (BV) and is now implemented in many breeding programs of different species such as e.g., cattle (e.g.: [20–22]), pigs (e.g.: [23, 24]), poultry (e.g.: [25, 26]), and aquaculture (e.g.: [27, 28]). In Switzerland, SNP genotyping data for genomic selection is only available in the economically most important cattle populations; however, the extent of genotyped animals varies greatly (Table 1).

1.2 Inbreeding

Inbreeding describes the process of mating related individuals. Especially in small and closed populations without any genetic influence from other populations inbreeding is inevitable, leads to a decrease in genetic diversity and will increase over time [17]. On a molecular level, inbreeding describes the probability of two alleles being inherited from the same ancestor [17]. Negative effects of inbreeding lead to the so-called inbreeding depression that has been observed for decades (e.g.: [29–33]). In a comprehensive study in Italian Holstein cattle, inbreeding depression was shown to affect production, reproduction, and functional traits [30, 31]. Doekes *et al.* (2019) differentiated between recent and ancient inbreeding, where recent inbreeding is assumed to be more harmful as artificial selection pressure can increase recessive alleles, while natural selection reduces them over time. This analysis of Dutch Holstein confirmed the negative effects of recent inbreeding on production, fertility, and health traits [31]. Traditionally, inbreeding had been estimated based on pedigree relationships; however, today inbreeding can be estimated by using genotype-based measurements, such as runs of homozygosity (ROH) (see chapter 1.4.1.2 Homozygosity mapping). Molecular genetic analyzes showed that inbreeding levels of small populations are comparable with large international populations [11, 34, 35]. It is hypothesized that in local breeds more natural service sires are used and therefore a greater number of sires per generation impact the genomic structure of these populations [11]. The excessive use of a few sires across the globe enhances the genetic gain but reduces effective population size. This trend was increased with the introduction of artificial insemination (AI) through which single bulls can be used intensively across the globe, as well as ovum pick-up and embryo transfer (*in vitro* fertilization) through which donor cows can have a manifold of offspring even at a young age by using recipients. In 2002, a survey showed that every fifth cow is bred by AI worldwide [36]. It can be assumed that this number grew in the last two decades. On the other hand, AI is used to produce crossbred offspring from dairy breeds, as the

cross would allow for the calves to go into meat production, as the necessity for replacement heifers has declined because of more productive dairy [37].

1.3 Reproductive success in cattle

From an agricultural point of view, maintaining production and genetic diversity within populations depends on the reproductive success of the living population [38]. Normally, once an animal is unproductive, e.g., an infertile cow, an outgrown pig, or a spent hen, it is removed from the population [39]. Interestingly, a decrease in longevity in livestock animals had been observed [39]. Even more, century-long artificial selection pressure on production traits left its marks, as can be seen by a decrease in fertility in dairy cattle [40, 41]. The unfavorable correlation between milk production and fertility is hypothesized to be the main force for this [40–42]. Generally, selection pressure on production traits and hitchhiking effects are the driving forces for reduced reproductive success [43]. A study regarding pregnancy loss in Simmental cattle exposed body condition score, milk yield, and parity as significant factors [44]. In addition, one of the major reasons for cattle to be slaughtered is lower fertility. Furthermore, belonging to certain breeds is also a risk factor for culling dairy [45, 46].

Hormonal imbalance, the health of the dam, and farm management can all affect the success of insemination. Additionally, viral, bacterial, and fungal infections can be detrimental during early embryogenesis [47–49]. Nevertheless, this thesis focuses on genetic causes affecting reproductive success.

Reproductive success depends on both the female and male sides. Within the scope of the herein presented studies, the focus is on the female perspective and recessive disorders segregating in various cattle populations. Nevertheless, there is a need in addressing the male perspective, which is carried out by colleagues of the Animal Genomics group at ETH in Zurich. In their studies, they analyze semen quality measurements of artificial insemination bulls, such as ejaculate volume, sperm concentration, sperm motility, sperm head and tail anomaly score, and the number of sperm filled per insemination straw [50–52]. They identified variants in the genes *WDR19*, *SPATA16, and QRICH2* on chromosomes 6, 1, and 19, respectively, associated with male fertility in BS bulls [50–52]. In international studies on Holstein cattle, dominant quantitative trait loci (QTL) on bovine chromosomes 8, 9, 13, 17, and 27 were found to be associated with sire conception rate [53, 54]. These QTL had been shown to harbor genes that have a fundamental role in the development of male reproductive organs, germline maintenance, and sperm maturation, including candidate missense variants in

the genes *AK9*, *TTLL9*, *TCHP,* and *FOXN4* [54, 55]. Fortes *et al.* (2013) summarized findings from association studies of various breeds (mostly Holstein and Angus). Traits regarding semen quality, development of reproductive organs, hormone concentrations, and effects on the birth process showed associations sporadically distributed over the entire genome [56].

1.3.1 Female fertility and rearing success

A major aspect of female fertility in cattle are ovulations that lead to multiple births [56]. At the University of Bern, a study on this subject is currently underway. Its first results identified a QTL on chromosome 11 between the genes *LHCGR* and *FSHR* in the Swiss Holstein population [57].

More important for the presented work is the identification of genomic regions associated with reduced fertility. Fertility, in general, can be further split into impaired fecundity, prenatal fetal loss, neonatal losses, and postnatal losses [56]. Impaired fecundity describes the incapability of a cow to produce offspring e.g., pregnancy loss or infertility. In breeding programs, fertility is often indirectly registered with traits linked to the number of inseminations and period between inseminations and births. From a genetic point of view, these could be caused by embryonic lethal variants, including protein-changing variants, aneuploidies, or variants with impact on the female estrus [49, 56, 58]. Many QTL affecting fertility have been published and are collected in the AnimalQTLdb [59]. With a focus on the Swiss populations, QTL associated with fertility and birthing traits were identified on chromosomes 13, 17, 19, 21, 25, and 29 in the BS population [60].

Reproduction success does not solely include successful inseminations and pregnancy losses, it also includes disturbances leading to impaired rearing success. This is even more important for beef breeders, where the calf is the main source of income. Unfortunately, there is only limited recording regarding the reasons why animals perish. Nevertheless, the Swiss recording system allows the registration of stillborn animals. Figure 2, using data from the recording system shows that the percentage of stillborn calves slightly increases over time and that there are more stillborn animals in dairy cattle than in beef cattle [9]. It can be assumed that these numbers are underestimated, as a recent study showed that perinatal mortality can range up to 28.2% on Swiss farms [61]. This study examined viral, bacterial, and fungal infections in both the calf and the dam, as well as morphological anomalies of the bodies and organs, that could explain roughly 34% of cases [61]. Furthermore, we assume that the peak in 2018 is caused by the

presented thesis project, as we actively collected cases of stillborn calves and thereby increased the recording of stillborn animals (Figure 2).

Figure 2: Percentage of stillborn animals in dairy (black) and beef (green) cows over the last decade (from [9]).

Besides stillborn animals, there is a multitude of dominant and recessive disorders known in cattle that have a lethal or sub-lethal effect. Interestingly, the first reports of recessive lethal disorders in cattle were published in 1928 several decades before the structure of DNA had been described by Watson and Crick in 1953 [62–64]. Ever since and especially with the inventions of second- and third-generation sequencing technologies allowing for rapid and cost-effective whole-genome sequencing (WGS), many Mendelian disorders have been described and associated with candidate causative variants. A great summary is a database called Online Mendelian Inheritance in Animals (OMIA) where currently 579 disorders in cattle are collected with potential causal variants described for 402 of them [65]. Furthermore, 296 of these phenotypes are potential models for human traits [65]. The presented thesis partially contributes to this collection of identified Mendelian disorders in cattle.

1.4 Forward genetic approach

In order to find genetic causes for disorders that appear in single cases or population-wide, two main approaches with a variety of methods are applied: forward and reverse genetic approaches. The **forward genetic approach (FGA)** uses phenotypic information of affected and non-affected individuals to associate the studied phenotype with the variation in the genome [66]. These case-control studies are best applied to monogenic Mendelian disorders, with a dominant or recessive mode of inheritance (MOI) (Figure 3A and D). The majority of fatal dominant gene defects can be attributed to *de novo* mutation events that occur in the parents (Figure 3E) or in the offspring (Figure 3F).

Caution needs to be taken if a disorder is located on the X chromosome and an X-linked inheritance pattern needs to be considered (Figure 3G, H, I, and J). Within X-linked disorders a sex bias is often observed, where only males or females are affected (Figure 3G and Figure 3J, respectively). More complicated are disorders with a multigenic or co-dominant inheritance, as the allele frequencies will not perfectly segregate with the phenotype (Figure 3C). Especially co-dominant inheritance is often complicated as the trait is inherited in a recessive manner so that all homozygous animals are affected, but the allele can also have complete and incomplete dominant effects and thereby lead to heterozygous affected animals (Figure 3C). Most of the methods used within FGA are described in detail below (see chapter 1.4.1 Selected methods of the forward genetic approach).

Within FGA, I investigated many dominant and recessive disorders, both autosomal and X-linked during the course of this thesis. I analyzed individual or multiple cases of skin disorders encompassing ectodermal dysplasia, ichthyosis, epidermolysis bullosa, Ehlers-Danlos syndrome, immunodeficiency with psoriasis-like skin alterations, and juvenile angiomatosis. The clinical picture of skin disorders can be extremely variable, based on genetic predisposition and numerous environmental effects such as housing conditions, feeding, and the presence of viral and/or bacterial pathogens [67, 68]. In general, a skin lesion carries a high risk of secondary infections and thus lowers the health and welfare of the animals [69]. Several genodermatoses in cattle have been previously described (Table A2) to be caused by small or large variants inherited in dominant, recessive, and X-linked modes of inheritance showing a vast heterogeneity of this group of diseases.

Furthermore, I investigated bone morphology affecting disorders, such as proportional dwarfism, chondrodysplasia, achondrogenesis, osteogenesis imperfecta, and hemifacial microsomia. These skeletal anomalies can have teratogenic, nutritional, or genetic origins [70]. Nevertheless, numerous genes had been shown before to be associated with these and similar disorders (Table A3). Some of these variants pose big issues for breeding, such as a variant in *ACAN* gene in the Dexter breed associated with disproportionate dwarfism or called bulldog calf syndrome in heterozygous or homozygous state, respectively [71].

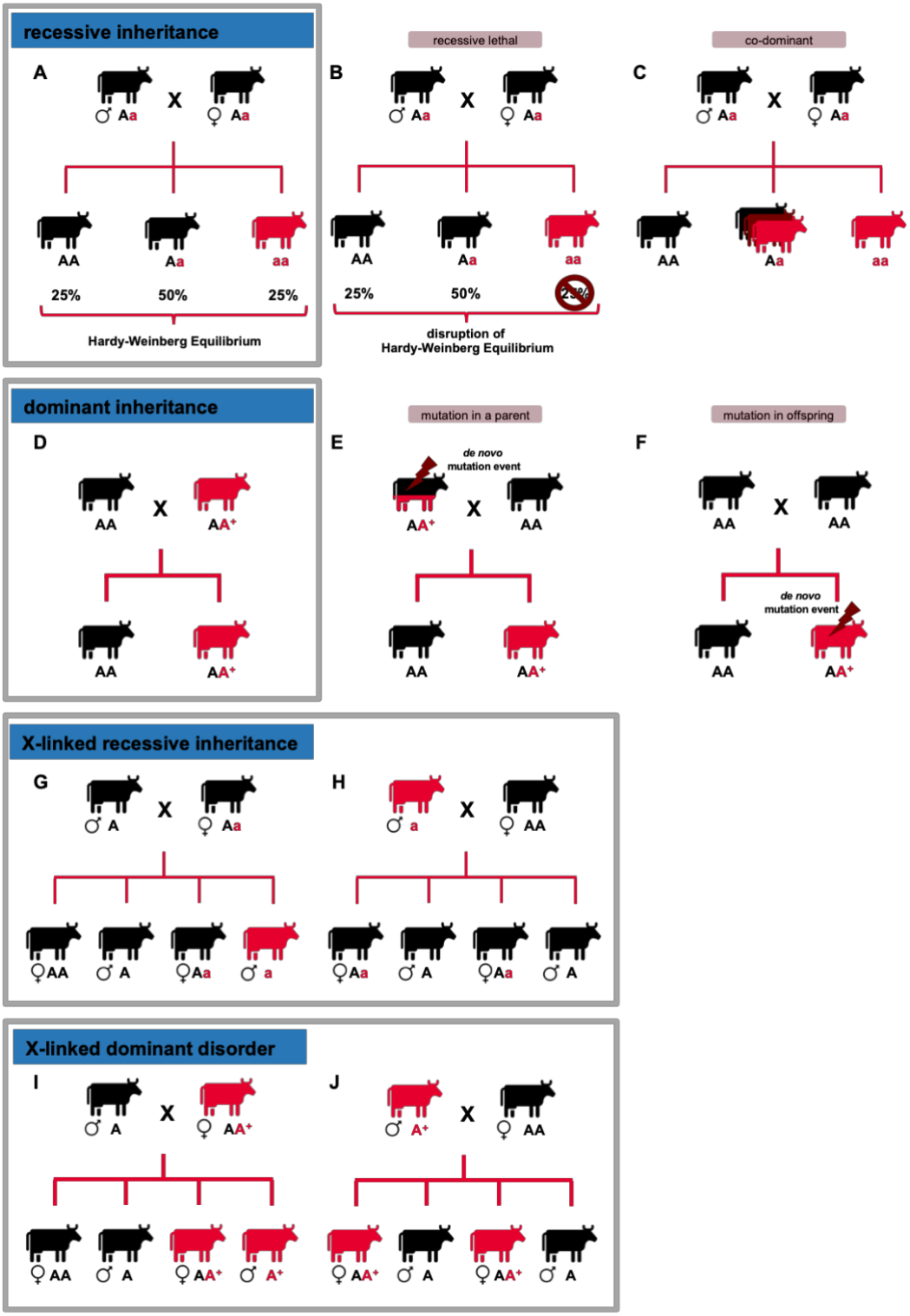


Figure 3: The most common modes of inheritance (MOI). The first block describes patterns of recessively inherited disorders: (A) the ordinary recessive inheritance, (B) the recessive lethal disorders, and (C) co-dominant disorders. Note that heterozygous animals of co-dominant disorders can be affected in a completely dominant manner, recessive manner, or any level in between. The second block described patterns of dominantly inherited disorders: (D) the ordinary dominant inheritance, (E) disorders caused by *de novo* mutation events in one of the parents, and (F) disorders caused by *de novo* mutation events in the offspring. The last two blocks describe X-linked recessive and dominant disorders. (G) and (H) illustrate the inheritance of X-linked recessive disorders, where the dam is a carrier of the causal variant (G), or the sire is a carrier of the causal variant (H). (I) and (J) illustrate the inheritance of X-linked dominant disorders, where the dam is a carrier of the causal variant (I), or the sire is a carrier of the causal variant (J).

I also investigated a neurological disorder in a crossbred calf showing neuromuscular channelopathy. Neuromuscular defects can be caused by exposure to toxic compounds, as well as due to neurogenic defects [72]. Known candidate variants associated with neurogenetic defects are presented in Table A4.

Furthermore, I analyzed more benign eye disorders causing achromatopsia and congenital cataract. In cattle, only strabismus, cataract, and progressive retinal degeneration have been described as inherited disorders of the eye, of which candidate variants were detected for the latter two (Table A5) [73–76]. In addition, some eye anomalies are described as symptoms of albinism in cattle [77, 78].

Lastly, during this thesis I examined fatal syndromic disorders that led to stillbirths showing severe anomalies such as pulmonary hypoplasia and anasarca (PHA) syndrome and skeletal-cardo-enteric dysplasia, as well as the effects of a known cholesterol deficiency disorder (Table A6) [79–82]. Table A6 also includes other known genes affecting reproductive and rearing success in cattle.

1.4.1 Selected methods of the forward genetic approach

Forward genetic approaches encompass any study, where phenotypes can be characterized. For inherited disorders, this usually includes case-control studies, where affected individuals (cases) are clearly distinguished from non-affected individuals (controls). Most of the herein described methods can be applied for both recessively and dominantly inherited disorders (Figure 4). All these approaches are most effective for monogenic disorders but can also be applied for the identification of QTL. The presented methods for population and family studies can be applied to SNP and WGS-derived single nucleotide variants (SNV) data. As results of genome-wide association studies (GWAS) are only presented in the results section of the reverse genetic approach, the method is further explained in section 1.5.1 Selected methods of the reverse genetic approach in chapter 1.5.1.2 Phenotype association studies.

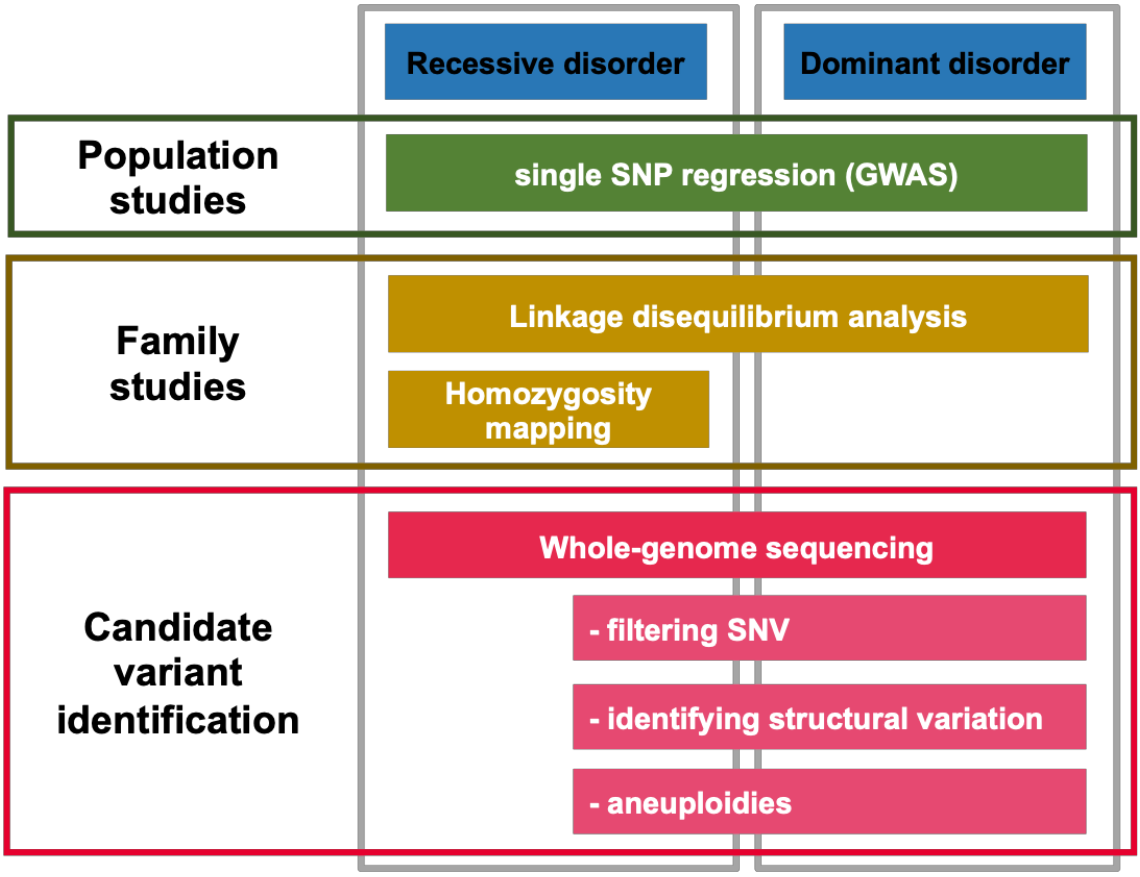


Figure 4: Potential analyzes included in the forward genetic approach. The green row represents population studies, brown represents family-based case-control studies, and red shows the possible use of whole-genome sequencing to identify candidate causative variants. If methods overlay with the left column, they can be used in recessive disorders and if they overlay with the right column, they can be used in dominant disorders.

1.4.1.1 Linkage analysis

Linkage analyzes (Figure 4) are applied to map genomic regions that are most likely causing a phenotype of interest in a family [83]. Therefore, the analysis is dependent on pedigree and genomic information of closely related individuals. With that information, the co-segregation of loci on the same chromosome can be estimated as the linkage disequilibrium (LD) that describes the non-random association of loci [83]. Within linkage analyzes, a parameter-based and parameter-free method can be applied [84]. While the parameter-based method requires an assumption of a mode of inheritance to analyze the co-segregation of marker and phenotype, the parameter-free method does not require any preliminary assumptions to analyze the probability of alleles to be identical by descent [83, 84].

1.4.1.2 Homozygosity mapping

Homozygosity mapping (Figure 4 and Figure 5) is most effective for recessively inherited disorders caused by a single mutation event. Therefore, it is expected that an affected animal carries two identical alleles [84]. Homozygosity mapping is also known as autozygosity mapping, as the method identifies alleles that are identical [85]. This is a powerful method to identify genomic loci that are shared among related affected individuals to locate alleles that are identical by descent (IBD). Whole sections of IBD alleles can be detected, as not only the disease-causing variant but also the flanking sequence is inherited together (Figure 5) [85]. Such stretches of homozygosity are called runs of homozygosity (ROH). The length of ROH correlates with the genetic distance, due to recombination events. Therefore, ROH are indicators for recent or ancient inbreeding and can be used to estimate inbreeding (see chapter 1.2 Inbreeding) [31].

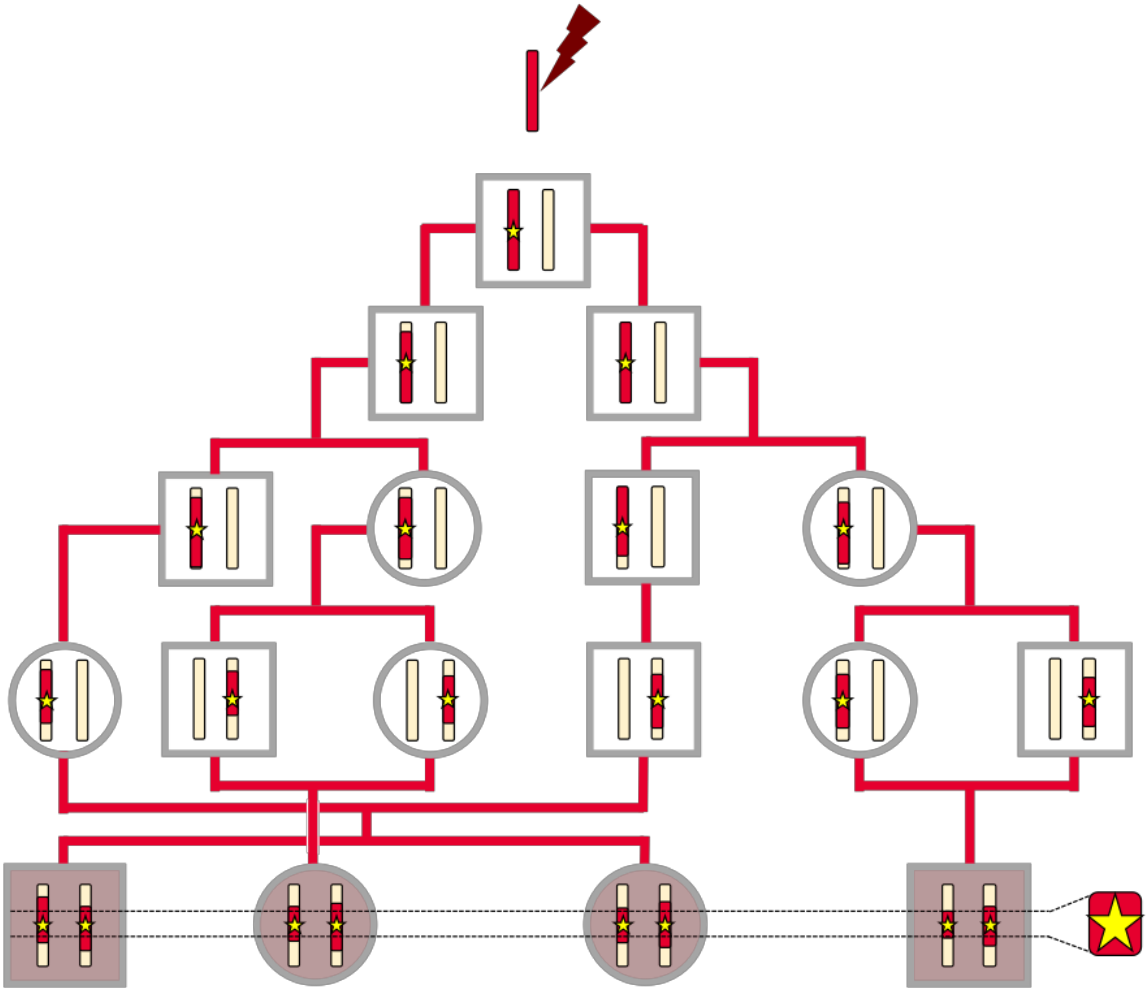


Figure 5: Illustration of homozygosity mapping. A recessive disease allele (yellow star) segregates in the population caused by a *de novo* mutation in the founder individual. Homozygous carriers are affected and carry the mutation and flanking sequence in the homozygous state (bottom row). Reproduced from [66].

1.4.1.3 Candidate variant identification

The identification of candidate causative variants for obvious or well-described phenotypes from WGS data is a multifaceted topic. For its success the identification of the right MOI and subsequent filtering of variants is vital (Figure 6). The most powerful approach for variant discovery in a case is to sequence the entire trio (dam, sire, and offspring); however, sequencing data of any additional related individual increases the probability of a successful detection [86]. As illustrated in Figure 6, recessive disorders need cases to be homozygous, their preceding generation to be heterozygous carriers of a candidate variant (obligate carriers), and sporadic carrier can be part of the unaffected population of controls. X-linked recessive disorders usually have male affected individuals that are predicted by bioinformatic algorithms to be homozygous carrier as a diploid organism is expected; however, only the dam is an obligate carrier (Figure 6). For dominant disorders that lead to the exclusion from the breeding population a *de novo* mutation event must have happened (Figure 6). Especially for *de novo* mutation events, it can be difficult to identify candidate variants, as it is often unclear at what stage a mutation event happened (Figure 7). Figure 7 illustrates that a *de novo* mutation can happen during embryogenesis and development of each parent or the offspring, or during the spermatogenesis or embryogenesis in the sire or dam, respectively [87]. Therefore, variants can be detected in blood, tissue or germline samples of the parents too (Figure 6 and Figure 7) [87]. Another possibility is that a parent is a chimera or is mosaic, carrying the variant already, but does not express symptoms due to low expression levels of the allele [87, 88]. Regarding dominant X-linked disorders, filtering is similar to autosomal dominant disorders, only that male cases are predicted to be homozygous carriers as with recessive X-linked variants (Figure 6). An important point to verify before filtering variants is if the affected animal is truly a single case, or if it might be a population issue that occurs sporadically. For the detection of candidate variants, narrowing down the genomic region through GWAS, ROH, or LD analyzes can help profoundly, especially if structural variants are involved. Furthermore, a candidate gene approach could be used; however, this is limited to disorders for which causative variants and associated genes have already been identified. This approach reinforces the research on non-human species, as they can become model organisms for human disorders [89].

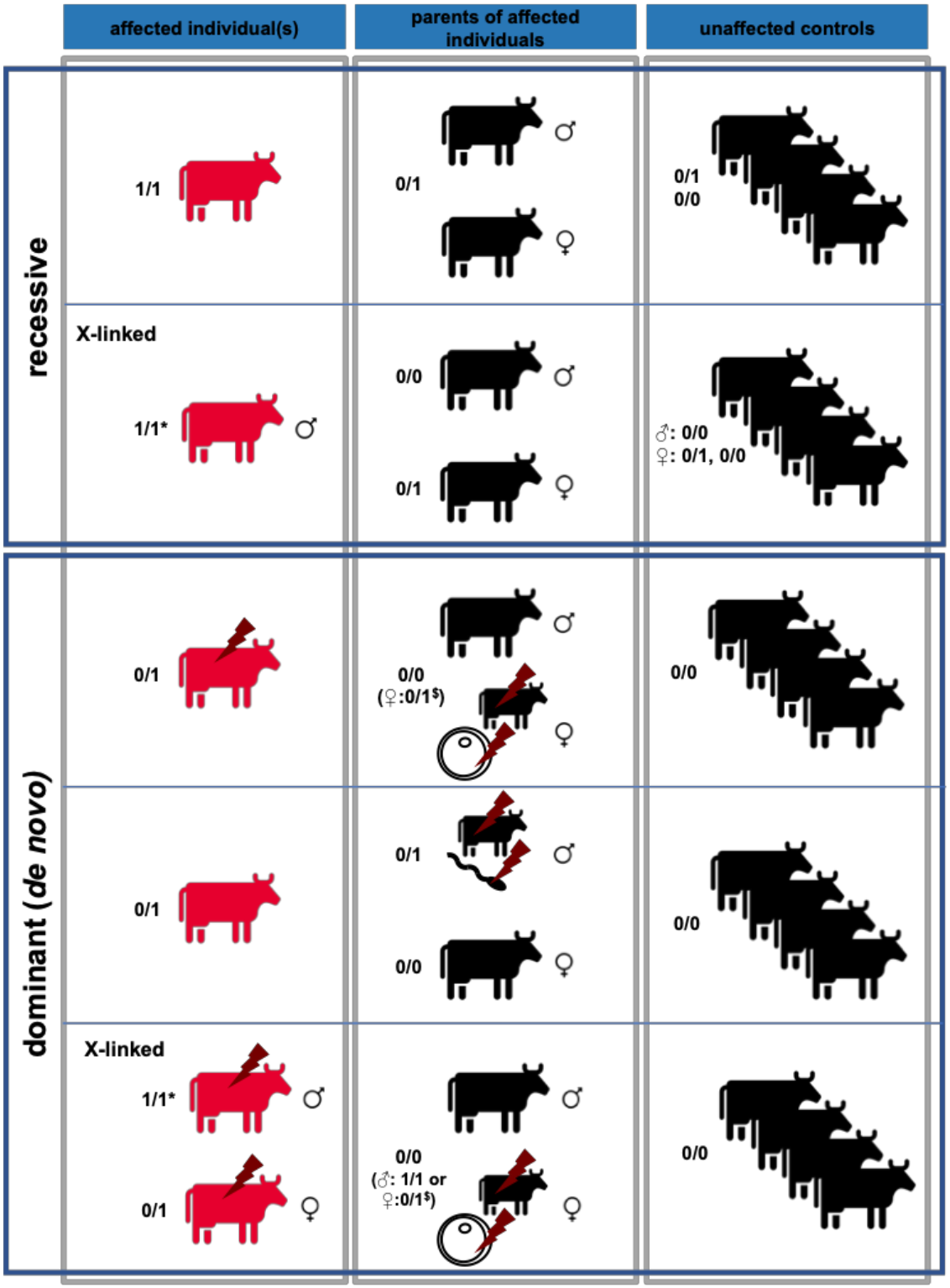


Figure 6: Systematic illustration on how to filter for candidate causal variants for obvious monogenic disorders in single nucleotide variants of the whole-genome sequencing data. Filtering depends on the mode of inheritance, on the underlying chromosome, and in case of *de novo* variants on the stage a mutation event happened (see Figure 7). Note that the prediction algorithm assumes diploidy and therefore predicts homozygosity for the haploid X-chromosome in males, indicated with an asterisk (*). The dollar sign ($) indicates potential *de novo* mutations in the female germline; however, this is hardly detected, as we usually do not have germline tissue of females available.

Regarding the type of variants, it is much simpler to screen the SNV catalog generated by bioinformatic analysis of the WGS data, than to identify structural variants (SV) (see chapter 1.6 Whole-genome sequencing). There is software available that helps filter through these massive files, such as VCFTools [90]. During the work on this thesis, a program to filter SNV for our specific projects was needed, whereby the program filterVariants evolved. This is a freely available bash script that runs on any Linux server (accessible on Github: ihaefliger/program_filterVariants [91]). The identification of SV is more complicated even if they are hypothesized to have a bigger impact on phenotypes than SNV [92]. More and more SV are being investigated based on whole-genome sequences and reported also in cattle [93–98]. These are mostly copy number variants (CNV), which describe the numerical differences of the copies of a part of the genome, including large deletions, duplications, insertions, translocations, and inversions. It is assumed that the number of copies directly affects the gene dosage. Not to forget that SV are estimated based on WGS data to cover roughly 16% of the bovine genome [94], while SNV of the currently available high-density bovine SNP array with 777k markers represent only about 0.03% of the genome. Nevertheless, the systematic detection of SV is still limited as detection algorithms seem to vary in performance depending on the type and size of structural variation [99]. Therefore, we often rely on the identification of candidate regions and the manual inspection of the sequencing data. For large deletions and duplications (several Mb), there were efforts done in our research group to produce scripts that plot the read depth across entire chromosomes and the genome (unpublished bash and R scripts: multicov). For smaller deletions and duplications (several kb), I implemented scripts to visualize the read depth on parts of chromosomes during the work on this thesis (accessible on Github: ihaefliger/program_coveragePlot_region [100]). A successful example applying these methods is the use of SNP data to identify a chromosomal translocation in piglets with a syndromic form of cleft palate, which was balanced in the sire and unbalanced in the affected offspring. The available WGS data made it possible to determine the exact breakpoints in the DNA sequence [101]. Nevertheless, for complex structural rearrangements, long-read sequencing outperforms short-read sequencing technologies as shown e.g., in a study of polled intersex syndrome variants in goats [102].

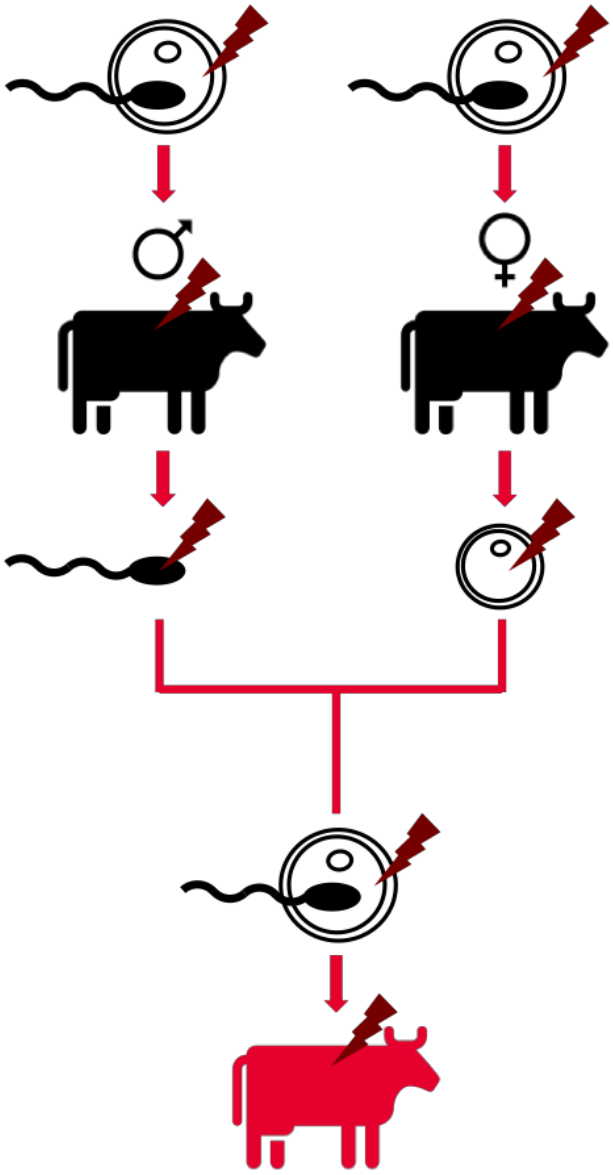


Figure 7: Illustration depicting the potential stages a *de novo* mutation can arise. Note, that the mutation event is indicated with the dark red arrow. A *de novo* mutation can arise during embryogenesis or development in both parents and lead to a mosaic individual. Further, each of the parent could have a *de novo* mutation during spermatogenesis or oogenesis. To lead to an affected individual after fertilization, a *de novo mutation* event can happen during embryogenesis or development in the offspring.

1.5 Reverse genetic approach

The second approach is the **reverse genetic approach (RGA)**, where population-wide massive genomic data is being screened for statistically significant deviations from basic genetic models indicating the segregation of recessive deleterious variants [66, 103]. The fundamental assumption is that the animals used in a breeding scheme are all healthy and could mate randomly [103]. If a recessive lethal variant is segregating and is in linkage disequilibrium with one of the SNV, these co-inherited SNV and their associated haplotypes will show a depletion in homozygous carriers (Figure 3C). In cattle, this can indicate a broad range of hidden phenotypes, such as early embryonic death, abortions during gestation, stillbirths, or weak calves that are non-viable or ill-thrifting. Most of the methods used within RGA are described in detail below (see chapter 1.5.1 Selected methods of the reverse genetic approach).

The work of VanRaden *et al.* (2011) represents the pioneering work proposing such an approach in cattle based on the increasing availability of massive SNP genotyping data used in breeding programs, in which they screen this data for significant depletions of the Hardy-Weinberg equilibrium (HWE). This kind of analysis has been used since 2011 for both beef [104–106] and dairy cattle [105, 107–111], as well as for pigs [112–114], chicken [115], and most recently horses [116], as it is vital for all livestock to have good reproductive capabilities. Most of the cattle analyzes identified numerous haplotypes with reduced homozygosity, that deviate significantly from HWE and are scattered across the bovine genome. For several of these haplotypes candidate causal variants were proposed (Table 2). Most variants have been associated with early embryonic death (HH1, HH2, HH6, HH7, JH1) or abortion (AH2, HH3, HH4, HH5, MH1, MH2, NH7); however, some have also been linked with increased mortality (AH1) and ill-thrift (BH2, FH2).

Table 2: Identified candidate causal variants associated with haplotypes showing depletion in homozygosity in different cattle breeds.

phenotype	haplotype	gene	OMIA	type of variant	breed	ref
Ptosis, intellectual disability, retarded growth and mortality (PIRM) syndrome	AH1	*UBE3B*	001934-9913	SNV (splice site)	Finnish Ayrshire	[117]
Abortion	AH2	*RPAP2*	002134-9913	SNV (splice site)	Ayrshire	[118]
Juvenile mortality	BH2	*TUBD1*	001939-9913	SNV (missense)	Brown Swiss; Fleckvieh	[103, 119]
Fanconi-Bickel syndrome	FH2	*SLC2A2*	001958-9913; 000366-9913	SNV (missense)	Fleckvieh; Original Braunvieh	[110, 120]
Abortion	FH4	*SUGT1*	001960-9913	SNV (missense)	Fleckvieh	[110]
Embryonic lethality	HH1	*APAF1*	000001-9913	SNV (nonsense)	Holstein	[103, 109, 121]
Embryonic lethality	HH2	*IFT80*	001823-9913	SNV (frameshift)	Holstein	[103, 122, 123]
Abortion	HH3	*SMC2*	001824-9913	SNV (missense)	Holstein	[103, 107, 108, 123]
Abortion	HH4	*GART*	001826-9913	SNV (missense)	Holstein	[109]
Abortion	HH5	*TFB1M*	001941-9913	gross deletion	Holstein	[81]
Embryonic lethality	HH6	*SDE2*	002149-9913	SNV (start lost)	Holstein	[124]
Embryonic lethality	HH7	*CENPU*	001830-9913	5bp deletion (splice site)	Holstein	[125]
Embryonic loss	JH1	*CWC15*	001697-9913	SNV (nonsense)	Jersey	[103, 126]
Abortion	MH1	*PFAS*	001827-9913	SNV (missense)	Montbeliarde, Vorderwald	[127]
Abortion	MH2	*SLC37A2*	001828-9913	SNV (nonsense)	Montbeliarde, Vorderwald	[109]
Abortion	NH7	*CAD*	002201-9913	SNV (missense)	Normande	[128]
Abortion		*RABGGTB*	002037-9913	SNV (missense)	Holstein	[129]
Abortion		*RNF20*	002038-9913	SNV (nonsense)	Holstein	[129]
Abortion		*TTF1*	002036-9913	SNV (nonsense)	Holstein	[129]

ref = references

1.5.1 Selected methods of the reverse genetic approach

The workflow applied within the presented thesis for the RGA is summarized in Figure 8. In the scope of this project, we screened the routine genotyping data of the breeds BS, OB, HO, and SI that were made available from the breeding associations (Table 3; Figure 8A). We applied two trio-based approaches, where the complete trio (sire, dam, and offspring) are genotyped (trio) and where the paternal half-sib group (sire, maternal grandfather, and offspring) are genotyped (pgp). The applied screens base the statistical analysis on exact tests of HWE, which allows it to be applied in small populations [130]. In all four breeds, it can be observed that the dataset of the pgp approach includes more groups than of the trio approach (Table 3). Nevertheless, the trio approach is very powerful, as it allows to trace the direct inheritance of the haplotypes.

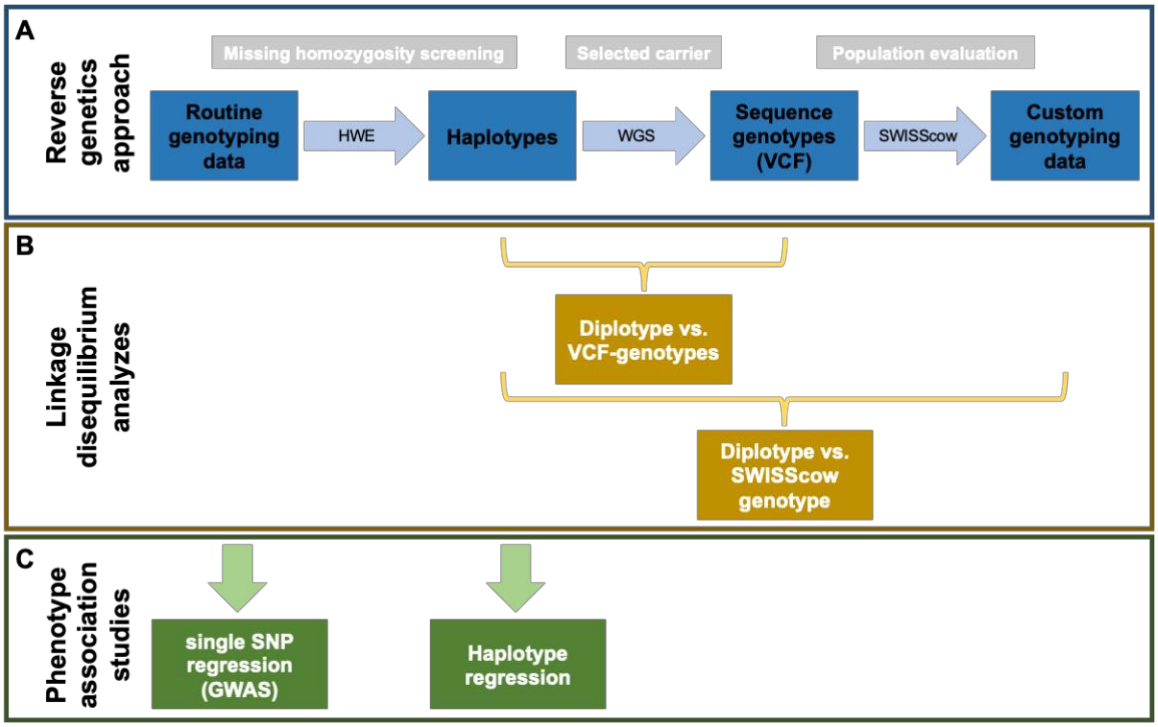


Figure 8: Flowchart of the reverse genetic approach. Depending on the data availability, in addition to the sole association of genomic information with a reverse genetic approach (A) and linkage disequilibrium (B), association analyzes can indicate phenotype effects (C).

Table 3: Number of genotyped animals and the data encompassed in the two approaches, pgp and trio, across the four breeds analyzed within the reverse genetic approach. Note that the numbers can differ from the latest numbers in Table 1.

	Original Braunvieh	Brown Swiss	Holstein	Simmental
number of genotyped animals	10,085	48,807	52,961	9,965
number trios (trio: sire, dam & offspring)	3,287	14,450	17,915	2,626
number paternal half-sib groups (pgp: sire, maternal grandsire & offspring)	4,360	32,319	30,315	3,969

Based on the output of the SNP genotyping data analysis, the identified haplotype carriers were selected in order to retrieve the WGS data of the individuals (Figure 8A). By whole-genome or -exome sequencing of carrier animals, one can screen a haplotype region for possible lethal variants and compare them to other healthy animals, which are not supposed to be homozygous carriers. Thereby, several causative variants have been previously identified (Table 2) [81, 104, 110, 119, 123, 124, 128, 131, 132]. Another approach was implemented in a study conducted in Belgium, that fully relinquished SNP array data, but used the genome-wide next-generation sequencing population data (Table 2) [129]. In their study, they estimated that between 11 and 17 recessive embryonic lethal protein-changing variants segregate in a population with an effective population size of 100, and each individual carries between 0.53 and 0.85, respectively [129]. Analogous to this, it was found that populations with low diversity carry more deleterious variants than populations with a high genetic diversity [133].

In the scope of my thesis, 130 animals distributed over the four populations of interest were selected to have their entire genome analyzed because they carried at least one haplotype. The WGS data generated was prepared according to the description in the following chapter 1.6 Whole-genome sequencing, and led to a comprehensive variant catalog, including the 130 animals and several hundred freely available control genomes from various breeds (Figure 8A). Further, a custom algorithm was implemented to filter for candidate causative variants. Basic restrictions on animals to be included in the dataset were a known health history without any disease records and sequenced at an average read depth of ≥6.8x. The genomic regions of interest were defined as the most significant haplotype plus flanking regions of ±2 Mb. Quality filters made sure that a marker showed a maximum of 50% missing genotype calls, and quality indication needed to be PASS (see chapter 1.6 Whole-genome sequencing for details). Further criteria on marker levels were to filter for SNV that occurred in an animal of the population of interest, never occurred in the homozygous state across populations, and showed a maximum carrier frequency of 75%.

For validation of the identified candidate causative variants, a big part of my project involved the design of a custom Affymetrix SNP array called “SWISScow” (Thermo Fisher Scientific, Waltham, Massachusetts, USA) (Figure 8A). This array was designed in collaboration with the Swiss breeding associations and includes SNP used in the routine genotyping, markers that arose through my RGA and FGA analyzes, previously known disease-causing variants, and additional protein-changing variants. The new design has been thoroughly tested by genotyping 192 samples with the SWISScow array twice and with the routinely used GeneSeek GGP Bovine 150k array from Neogene (Neogene, Santa Monica, California, USA), including closely related animals (trios), carriers of known disease-causing variants, and animals with WGS data available. With this data, the repeatability of the SWISScow array was tested and the concordance of the SWISScow array and the routine genotyping array, as well as the sequencing data, was verified. Furthermore, our SWISScow array was evaluated to work accurately in the context of gender recognition, verification, and detection of family relationships, as well as the detection of known genetic disorders. Finally, the SWISScow array included 112,854 routinely used SNP and 205,362 other variants. For a year now, the SWISScow array has replaced the previously used SNP array, which allowed for the genotyping of the current Swiss cattle breeding populations. Thereby, 13,667 cattle were genotyped with the SWISScow array including 6,575, 1,489, and 5,603 from the BS, OB, and HO populations, respectively. In addition, 574 SI animals had been genotyped with the SWISScow array, which is not enough data to gain strong statistical power.

Finally, candidate causative variants could be proposed based on the described analysis purely on genomic information and statistical evidence. However, solely protein-coding variants were proposed as candidate causative variants because variant effect predictions can be performed for them more easily.

1.5.1.1 Linkage analyses

To make sure that the identified candidate causative variants are associated with the detected haplotype, linkage analyzes were performed (Figure 8B). Mostly, LD analyzes are performed to identify genomic regions that are inherited together and thereby associated with a trait of interest segregating in closely related families [83]. Contrary to this, in RGA we used LD as a measure to verify if a variant is inherited together with the associated haplotype. As it was assumed that a mutation causes a haplotype to show reduced homozygosity, this mutation would have to be co-inherited with the haplotype at least up to a certain level. Therefore, LD was calculated with plink v1.9 [134] based on the diplotype states of the haplotype and the SNP on the SWISScow array, as well as

the WGS-derived SNV. While the analysis with the SWISScow data could be done on all the individuals genotyped, the WGS data was more sporadic, as only a few animals per population had WGS data available.

1.5.1.2 Phenotype association studies

Lastly, to get a feeling for the effect of the identified genomic regions harboring recessive variants, phenotype association studies were performed if there was enough data available (Figure 8C). Phenotype information included pseudo-phenotypes represented by deregressed breeding values from the routine breeding value estimation of a broad range of fertility, birth, growth-related, and survival traits (Table A1) [135].

Genome-wide association studies (GWAS) are commonly performed based on SNP data; however, genotypes derived from WGS data can be used as well [66]. A GWAS analysis allows finding markers that segregate together with a trait based on their allele frequency. A common example are case-control studies, where a marker that always occurs in cases and only sporadically in controls would lead to a significant association. GWAS helps to identify loci in populations for monogenic traits as well as for quantitative traits [136]. Mathematically, a GWAS is a single SNP regression based on a model such as: $y = \mu + G + \mathbf{\beta S} + \epsilon$, where y represents the phenotype, μ is the average phenotype, G is a random effect that includes a pair-wise genomic relationship matrix, $\mathbf{S}$ is a vector with the genotypes per animal, $\mathbf{\beta}$ is the estimated fixed effect of each genotype and ϵ accounts for the random variation. Due to the repeated test of the model, as each SNP individually is tested for an association, it is important to correct for multiple testing to avoid sporadic associations e.g., using false discovery rates [137], or Bonferroni correction [138]. Furthermore, it is important to understand that associated SNP are not necessarily causative but merely linked to a causative marker within genomic regions of interest. The most common visualization of GWAS results is the so-called Manhattan plots that can be created using R with the package qqman [139, 140].

On one hand, GWAS results were used to verify if identified haplotypes co-segregate with QTL of the Swiss populations. On the other hand, the effects of the haplotypes on the breeding values could be estimated in a model analogous to the GWAS, but with only one marker representing the haplotype. These association studies were performed to get an indication of the nature of a potential hidden phenotype.

1.6 Whole-genome sequencing

Only in the last decade has next-generation sequencing made it possible to unravel an individual's entire genomic DNA in a cost- and time-efficient manner. WGS allows comprehensive insights into the organization and functionality of the genome.

1.6.1 Samples

For my herein presented studies, samples used for DNA analysis were taken from blood (EDTA), semen, hair roots, or tissue. Regarding EDTA-blood, the Nucleon Bacc2 kit (GE Healthcare) was used to extract the DNA. QIAGEN's DNeasy kit was used for semen, hair roots, or tissue samples following the manufacturers' instructions. The Maxwell RSC instrument (Promega) was used to perform DNA extraction. Generally, PCR-free fragment libraries are prepared to perform paired-end sequencing of 150bp reads and aim at a read depth of 20x. At the start of this PhD project, the sequencing platform Illumina HiSeq3000 was used, later from 2018 onward, the sequencing platform NovaSeq was used. Older WGS data would also have been produced on the platforms HiSeq2000 and HiSeq2500. All WGS data produced for the purpose of the presented studies were transferred to the public repository European Nucleotide Archive (ENA) as part of the ongoing Swiss Comparative Bovine Resequencing project (Accession ID: PRJEB18113) [141].

1.6.2 Workflow

As the WGS data is an integral part of the variant identification of the presented studies, the workflow applied to the raw data shall be introduced (Figure 9). As the *de novo* assembling of a genome is computationally highly complex as well as a time- and memory-consuming task, a reference sequence and an alignment approach for the sequenced reads are predominantly used. Therefore, we can detect any kind of variation of an individual sequenced genome to the reference sequence. In order to do so, cattle geneticists use an analogous pipeline to the GATK best practices mostly applied in other species [142], which is adjusted to the needs of cattle genomes. Furthermore, the preparation is in accordance with the recommendations of the 1000 Bull Genomes Project (run8) [143], as the data was submitted to the data collection of the consortium. The consortium published their first dataset in 2014 with 234 whole-genome sequenced cattle [144]. Their latest dataset of run9 in 2021 encompassed 5,116 genomes from across the globe, distributed over 100 breeds and diverse crosses. The most prominent

breeds are Holstein, Angus, Norwegian Red, and Brown Swiss with 1'148, 401, 347, and 294 purebred individuals, respectively.

The workflow starts with the raw data arriving in the fastq file format. These files were quality controlled by trimming reads with fastp v0.12.5 and the settings: *--qualified_quality_phred 20, --length_required 35, --cut_window_size 3, --cut_mean_quality 15, --cut_by_quality5 20, --cut_by_quality3 20* and *--trim_poly_g* [145]. Next, the quality was checked batch-wise with the programs FastQC v0.11.5 [146] and MultiqC v1.8 [147] to see if any files showed reduced quality across the sequences.

The quality-controlled data was then mapped using the Burrows-Wheeler Aligner v0.7.17 [148] to the cattle reference sequence published in 2018 ARS-UCD1.2 (see chapter 1.6.3 Reference assembly) [149, 150]. If there were several bam-files per animal, the files were merged using the command *MergeSamFiles* from the software picard-tools v.2.18.2 [151] and sorted with the command *sort* of the software samtools v1.8 [152]. Furthermore, the bam-files were controlled for duplicate reads with the command *MarkDuplicates* from the software picard-tools v.2.18.2 [151] and recalibrated using the commands *BaseRecalibrator* and *PrintReads* of the software GATK v3.8.1.0.gf15c1c3ef [153]. Recalibration was performed on the variants known from the 1000 Bull Genomes Project run 7 (BQSR file version 2) [143]. Finally, the bam-file was indexed with the command *index* of the software samtools v1.8 [152]. These bam-files can be visualized in the Integrative Genomics Viewer (IGV) [154] to inspect candidate variants.

From the bam-files, genomic variant call format (VCF) files per animal were produced including small insertions and deletions (InDel), as well as SNV. These files encompass all positions in the genome that differ from the reference sequence and were produced with the command *HaplotypeCaller* from the software GATK v3.8.1.0.gf15c1c3ef [153]. After that, all the single genomic VCF files are combined into a single VCF file including all variants from all individuals by using the commands *CombineGVCFs*, *CatVariants,* and *GenotypeGVCFs* from GATK v3.8.1.0.gf15c1c3ef [153].

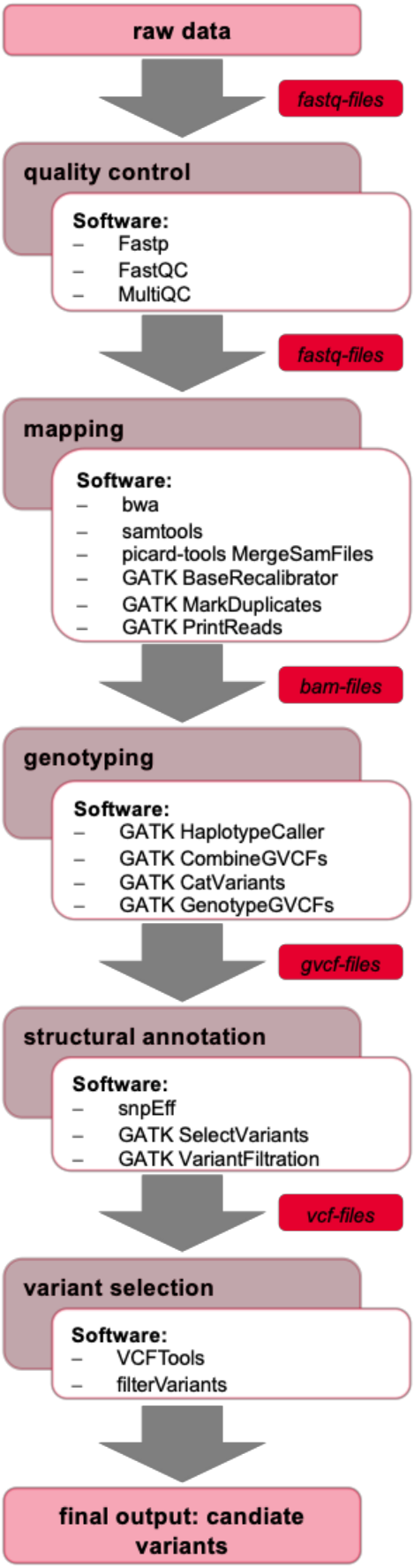


Figure 9: Example of the bioinformatics workflow applied for mapping of whole-genome sequencing data to a reference sequence and calling variants.

For every single variant included in the comprehensive VCF, a quality indicator and the effect of the variant on the encoded protein were predicted. Regarding the effect prediction, the NCBI Annotation Release 106 [155] was used within the software SnpEff v4.3 [156]. In general, the effect prediction of SNV is not an easy task, especially if a variant affects non-coding sequence. Variants in the coding sequence can be predicted for their protein-changing impact rather simply, as they lead to either synonymous amino-acid exchanges denoted to have LOW impact, non-synonymous amino-acid exchanges denoted to have MODERATE impact, and frameshifts or nonsense variants that are denoted to have HIGH impact. The quality indicators were calculated by using the commands *SelectVariants* and *VariantFiltration* of GATK v3.8.1.0.gf15c1c3ef [153]. Thereby, hard filters are applied for InDel and SNV separately based on the GATK recommendations. These include for SNV to get the quality indication PASS to pass all the thresholds (*QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, and ReadPosRankSum < -8.0*) and to get the quality indication GATKstd the thresholds at lower level (*QD < 1.0, FS > 100.0, MQ < 30.0, MQRankSum < -15.0, and ReadPosRankSum < -12.0*). Any SNV not passing the thresholds gets the indication LowQual. Analogously, the PASS indication is given for InDel passing the thresholds for insertions and deletions (*QD < 2.0 and FS > 200.0*), GATKIndelStd is given if the reduced thresholds are passed (*QD < 1.0 and FS > 300.0*), and LowQualIndel for all other InDel.

1.6.3 Reference assembly

When applying the bioinformatics approach described above to individual whole-genome sequence data, the quality of the reference sequence determines the quality of variant detection that can be achieved. In 2009, the first genome draft of an inbred Hereford cow Dominette was published based on a 7x coverage (Figure 10) [157]. The DNA of Dominette would be the base for the first internationally available reference sequence Bos_taurus_UMD3.1.1 (GenBank: GCA_000003055.5; RefSeq: GCF_000003055.6) available in the database of the National Center for Biotechnology Information (NCBI) [158].

Figure 10: Dominette, inbred Hereford cow that the NCBI reference sequences are based on [159].

Currently, there are four cattle reference sequences registered in NCBI, of which three are based on the biological sample of Dominette (Figure 10; Table 4) and have an annotation available. The latest internationally used reference sequence is the ARS-UCD1.2 genome assembly (GenBank: GCA_002263795.2; RefSeq: GCF_002263795.1) that was published with annotation in 2018 in the NCBI database [149, 150, 155]. This reference sequence was used throughout my thesis.

Lately, novel breed-specific genome assemblies were published. Most of them used the trio binning approach, where allelic variation is introduced to allow assemblies based on haplotypes [160]. Therefore, F1 crosses between evolutionary distinct species or subspecies (e.g., *Bos taurus taurus* X *Bison taurus indicus*) are used to enable the separation of haplotypes and produce haplotype-based *de novo* assemblies. With this method novel genome assemblies were published for Angus [160], Brahman [160], Simmental [161], and Highland cattle [162], as well as for the species Bison [163] and Yak [162]. These trio-based reference sequences show high quality and even exceed the quality standard of the latest reference sequence (Table 4).

Table 4: Overview of the Hereford and Brown Swiss genome assemblies available in the database of the National Center for Biotechnology Information (NCBI) and the published Simmental genome assembly [161]. Note that the Simmental assembly is produced by trio-binning approach and the others by traditional *de novo* assembly.

	Hereford 2014[a]	**Hereford 2015[b]**	**Hereford 2018[c]**	**Brown Swiss 2021[d]**	**Simmental[e]**
Total sequence length (Gb)	2.670	2.725	2.716	2.658	2.862
Scaffold count	6,336	5,998	2,211	14,725	1,315
Scaffold N50 (Mb)	6.4	6.8	103.3	26.0	102.5
Scaffold L50	107	104	12	29	12
Number of contigs	75,617	42,267	2,597	34,351	1,315
Contig N50 (Mb)	0.097	0.276	25.9	0.268	70.8
Contig L50	7,930	2,849	32	2,856	14
Number of chromosomes	30	31	31	*NA*	30

[a]Bos_taurus_UMD3.1.1: GCA_000003055.5
[b]Btau_5.0.1: GCA_000003205.6
[c]ARS-UCD1.2: GCA_002263795.2; used in this thesis
[d]Brown Swiss: GCA_914753205.1
[e]ARS-Simm1.0 [161]

2 Hypothesis and aim

The general hypothesis of this thesis is that monogenic variants segregate in cattle and cause disorders with lethal, sub-lethal, or damaging impact on hidden as well as visible phenotypes leading to exclusion from the breeding population.

For obvious disorders that can be described and sampled, a straightforward identification of potential candidate variants was performed. In such cases, the **forward genetic approach** including whole-genome sequencing was applied to various affected cattle to describe the genetic etiology of the underlying disorder. The described disorders are of various modes of inheritance and represent individual cases of various breeds, as well as population issues.

Animals affected by the following group of disorders were analyzed:

- Inherited disorders affecting the skin (genodermatoses)
- Inherited disorders affecting bone morphology
- Neuromuscular disorders
- Heritable eye disorders
- Fatal syndromic disorders (congenital anomalies)

For hidden disorders, no direct sampling can be undertaken. Therefore, the second part of this thesis aimed to apply a **reverse genetic approach** to exploit massive genome-wide genotyping and sequencing data to identify haplotypes and variants causing hidden phenotypes. To validate the obtained results, statistical association with routine phenotypes, as well as population-wide genotyping was applied.

This approach was performed in four Swiss dairy populations:

- Brown Swiss
- Original Braunvieh
- Holstein
- Original Simmental

3 Results

3.1 Results table of contents

3.2 Results forward genetic approach

An *IL17RA* frameshift variant in a Holstein cattle family with psoriasis-like skin alterations and immunodeficiency

Journal:	BMC Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, genetic figures, writing original draft and revisions
Displayed version:	Published version
DOI:	10.1186/s12863-020-00860-4

Häfliger et al. BMC Genetics (2020) 21:55
https://doi.org/10.1186/s12863-020-00860-4

BMC Genetics

RESEARCH ARTICLE Open Access

An *IL17RA* frameshift variant in a Holstein cattle family with psoriasis-like skin alterations and immunodeficiency

Irene M. Häfliger[1†], Marlene Sickinger[2†], Mark Holsteg[3], Leif M. Raeder[4], Manfred Henrich[4], Siegfried Marquardt[5], Cord Drögemüller[1*†] and Gesine Lühken[6†]

Abstract

Background: Skin lesions and dermatoses in cattle are often associated with infections due to bacteria, fungi or environmental risk factors. Dermatoses with genetic etiology have been described in cattle. Among these rare disorders, there are primary congenital dermatoses that are associated with inherited nutritional deficiencies, such as bovine hereditary zinc deficiency or zinc deficiency-like syndrome. This study presents three cases of Holstein cattle with congenital skin lesions observed on a single farm that resemble zinc deficiency-like syndrome. Close clinical and pathological examinations took place in two cases. Pedigree analysis indicated autosomal recessive inheritance and whole-genome sequencing of both affected calves was performed.

Results: The two calves showed retarded growth and suffered from severe ulcerative dermatitis with hyperkeratosis, alopecia furunculosis and subcutaneous abscess formation. Blood analysis showed correspondent leukocytosis with neutrophilia whereas minerals, macro- and micronutrients were within the reference ranges. Variant calling and filtering against the 1000 Bull Genomes variant catalogue resulted in the detection of a single homozygous protein-changing variant exclusively present in both sequenced genomes. This single-nucleotide deletion in exon 3 of *IL17RA* on bovine chromosome 5 was predicted to have a deleterious impact on the encoded protein due to a frameshift leading to a truncated gene product. Genotyping of the affected cattle family confirmed recessive inheritance.

Conclusions: A loss-of-function mutation of the IL17RA transmembrane protein could be identified as most likely pathogenic variant for the psoriasis-like skin alterations observed in the two affected Holstein calves. In man, rare recessive diseases associated with *IL17RA* include immunodeficiency 51 and chronic mucocutaneous candidiasis. This supports the observed immunodeficiency of the presented cases. This study reports the first naturally occurring *IL17RA*-associated animal model.

Keywords: Cattle, Genetic disorder, Monogenic, Mendelian, Skin disorder, Precision medicine, Rare disease, Interleukin 17 receptor a

* Correspondence: cord.droegemueller@vetsuisse.unibe.ch
†Irene M. Häfliger, Marlene Sickinger, Cord Drögemüller and Gesine Lühken contributed equally to this work.
[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland
Full list of author information is available at the end of the article

Background

An intact coat and skin is of utmost importance for the protection of the body from environmental impacts, including infections and thermal challenges. Clinical dermatological conditions vary in their dimensions, from alopecia, hypo- and hyperkeratosis, coat discoloration, seborrhoea to pruritus. Often several symptoms occur simultaneously. Dermatoses are reported frequently and can be due to a variety of causes. The aetiology of most skin disorders is based on a combination of genetic and environmental factors representing multifactorial or complex diseases. Nevertheless, some are determined exclusively by the environment, e.g. due to infections with pathogens belonging to bacteria, viruses, fungi and parasites [1–4], intoxication [5] and dysfunctional housing [6]. Skin lesions, e.g. due to injuries are at great risk for secondary infections. Generally, they can lead to a reduced animal welfare, performance and therefore to a loss of the economic value of an animal. In cattle, the most common and frequently endemic skin disorder in Central Europe is due to infection with dermatophytes [7]. More rarely, dermatoses occur that are determined exclusively by genetic factors, such as e.g. epidermolysis bullosa [8], or that are associated with inherited disorders leading to nutritional deficiencies of e.g. vitamin A and C, zinc, copper, or fatty acids due to malabsorption [9]. Genetic diseases of the skin are called genodermatoses and typically follow a monogenic mode of inheritance, which means that the genotype at a single gene determines whether the trait is expressed or not [10]. A hallmark of genodermatoses is familial clustering of cases and recently molecular veterinary genetics has made significant advances in the analysis of hereditary dermatoses [10].

A known genetic disorder segregating in the Holstein breed and leading to recurrent bacterial infections, delayed wound healing and retarded growth is the bovine leukocyte adhesion defect (BLAD; OMIA 000595–9913) [11, 12]. This defect is due to the p.Asp128Gly missense variant in the *ITGB2* gene [11]. The clinical symptoms induced by this variant are lesions and inflammations on oral mucous membranes and teeth, along with chronic pneumonia and diarrhoea [12].

Another autosomal monogenic recessive condition primarily seen in Holstein and Fleckvieh dairy cattle with psoriasis-like skin lesions in young calves was suspected to be the consequence of a secondary zinc deficiency [13]. It was assumed that a congenital gastrointestinal zinc malabsorption leads to an impaired function of the immune system, growth retardation and severe skin alterations (acrodermatitis). In Holstein cattle this disorder is known as bovine hereditary zinc deficiency (BHZD, OMIA 000593–9913) for a long time and reported to be associated with a recessively inherited splice-site variant in the bovine *SLC39A4* gene [14]. The altered SLC39A4 protein is described to lack two motifs, which lie in adjacent transmembrane domains involved in the construction of a pore responsible for the transport of zinc. Therefore the reported *SLC39A4* variant is most likely responsible for the impaired zinc absorption in this disorder [14]. The identification of another pathogenic variant associated with the so-called zinc deficiency-like (ZDL) syndrome (OMIA 001935–9913), a highly similar genodermatosis occurring in Fleckvieh cattle indicated obvious locus heterogeneity for this entity of disorders. Clinically a massive hyperkeratosis in combination with secondary microbial infections were observed in the ZDL-affected animals carrying two copies of the unravelled nonsense mutation in bovine *PLD4* [15]. In addition, all BHZD as well as ZDL cases were also underdeveloped in body size and weight and had a history of recurring diarrhoea and pneumonia [13–16]. Although there is a striking similarity in the phenotypic appearance of these two disorders, there are no clues for a functional connection between PLD4, an uncharacterized member of the family of phospholipid signalling enzymes, and zinc metabolism. As the ZDL-affected Fleckvieh calves did not respond to a dietary zinc supplementation, an impaired zinc metabolism is most likely not causing this disorder.

Herein we report the detailed phenotype of three half-sib purebred Holstein calves with severe congenital psoriasis-like skin alterations, retarded growth and an enhanced susceptibility to different infectious diseases such as bronchopneumonia. Clinical, genealogical, pathological and histopathological examination of the calves suggested a rare and recessive genodermatosis resembling ZDL. However, ZDL has not been observed in Holstein cattle until now. Our genetic analysis revealed a perfectly linked recessively inherited frameshift variant in the *interleukin 17 receptor A* (*IL17RA*) gene, which is most likely disease causing.

Results

Clinical findings

Initially, three calves were presented to the local veterinarian for clinical examination and therapy. One of the calves suffered from a severe bronchopneumonia and died shortly after presentation despite immediate therapy. The two other calves (3 and 5 months old) showed retarded growth and development even though they had been clinically inconspicuous within the first 4 to 6 weeks of their lives. Chronic skin lesions with signs of severe inflammation were present especially in skin folds, beneath the tail and at the umbilical region. Blood analysis including zinc supply, examinations for infections with bluetongue virus (BTV) or lumpy skin disease (LSD) resulted unsuspicious.

These two female calves were admitted to the Clinic for Ruminants (Internal medicine and Surgery) of the Justus-Liebig University Giessen for a deeper clinical investigation. The 5-months old calf will further be referred to as case 1 and the 3-months old calf as case 2. Both cases were lethargic and slightly dehydrated and showed massive hyperkeratotic alterations of the skin with regional alopecia and signs of inflammation. Skin lesions were partly bleeding and a sanious odour of the skin could be noticed (Fig. 1). In case 1, the head was asymmetric due to severe swelling in the region of the left mandibular lymph nodes (Fig. 1a). It was a 20 × 20 × 10 cm large mass in front of the shoulder joint. This calf also showed inflammations at the inguinal skin fold. Case 2 mainly showed marked inflammation of the axillar skin (Fig. 1f) and at the inguinal skin fold (Fig. 1b and e) as well as small abscesses at the lower surface of the tale (Fig. 1) and jaw (Fig. 1d). The palpable lymph nodes were markedly swollen in this calf.

Results of the peripheral blood analysis regarding the standard parameters are summarized in additional file 1 for each calf individually. The numbers of leukocytes, segmented neutrophils and the lymphocytes deviated from the normal range, indicating inflammation. The value for glutathione peroxidase indicated a sufficient supply of selenium in both calves. Regarding the minerals, solely magnesium showed a considerably reduced serum concentration, while zinc showed a mildly increased concentration.

Furthermore, the microbiological analysis of skin biopsies and lung tissue samples are provided in additional file 2 for both calves. Comparing the levels of infection, a high content of specific bacteria was more often observed in case 2 than in case 1, and comparing the tissue types, the bacterial content on skin was slightly higher than in the lungs. No fungal pathogens were determined in skin or lung samples. Tests concerning bovine virus diarrhoea (BVD), bluetongue virus (BTV) and bovine poxvirus (BPV) resulted negative.

Necropsy

Both calves showed similar gross lesions and a deferred development with a body weight of 113 kg (case 1) and 95 kg (case 2). The dermis revealed an ulcerative dermatitis and pododermatitis with pustules, alopecia, furunculosis, subcutaneous abscess formation, and marked

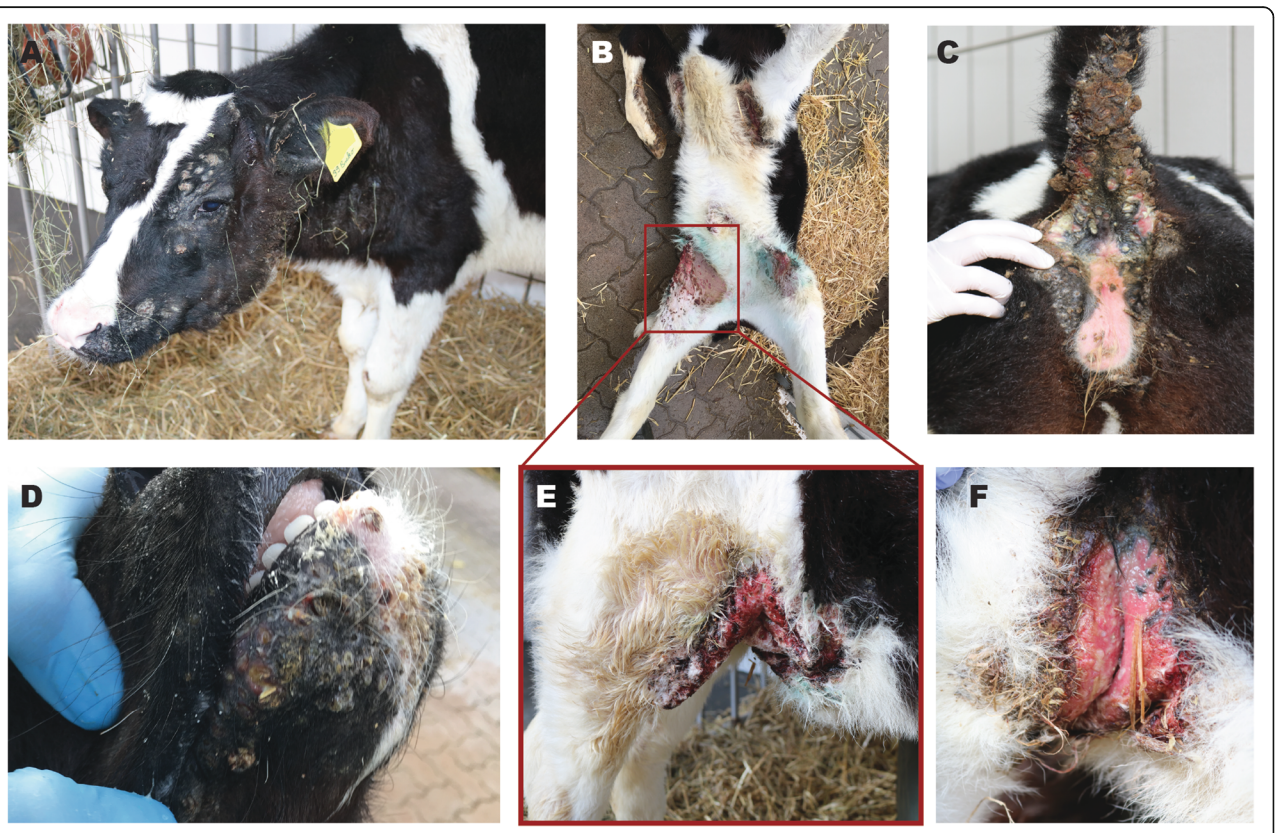


Fig. 1 Features of the skin anomalies in case 1 (A) and case 2 (B-F). **a** Head with massive swelling at the left mandibular angle. Palpation of this mass displayed fluctuation and severe phlegmon of the skin. **b** Ventral view on the lesions in the inguinal skin. **c** Abscesses beneath the tail. Crusts of pus and incrusted faeces are present. **d** Lesions around the mouth. **e** Right inguinal skin fold with massive inflammation and ulceration. **f** Axillar skin lesions. Palpation of the inflamed skin resulted in instant bleeding

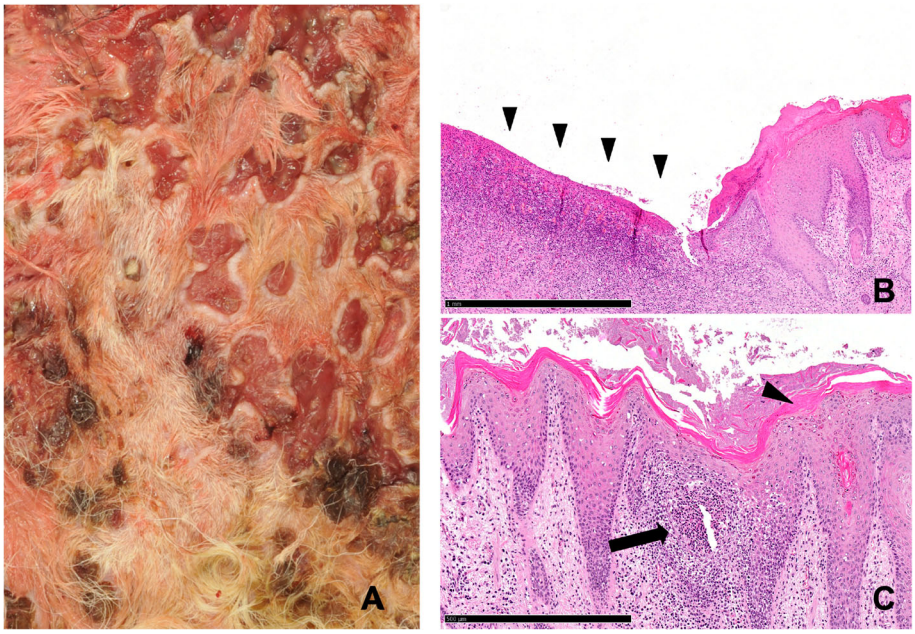


Fig. 2 Details of the inguinal skin lesions of case 2. **a** Macroscopic image of chronic ulcerations with serocellular crusts. **b** Microscopic view of severe ulceration of the epithelium (arrowheads) with formation of granulation tissue and necrosuppurative inflammation. Hematoxylin and Eosin, 5x, bar = 1 mm. **c** Microscopic view showing marked orthokeratotic hyperkeratosis (arrowhead). Hematoxylin and Eosin, 10x, bar = 500 μm

orthokeratotic hyperkeratosis (Fig. 1a-f). Mainly the areas of the axillar skin (Fig. 1f), the inguinal skin (Fig. 1b,e, Fig. 2a), the tale (Fig. 1c) and the mucocutaneous junctions were affected. Additionally, there was a severe epithelial hyperplasia and a moderate perivascular infiltration with lymphocytes, plasma cells, and macrophages (Fig. 2). Microscopically the massively inflamed skin lesions at the skin folds are visualized in Fig. 2b and c, where the chronic ulcerations with serocellular crusts are visible (Fig. 2a). The junctions show ulcerations of the epithelium with formed granulated tissue and the inflamed cells (Fig. 2b) and orthokeratotic hyperkeratosis with pustules were detected (Fig. 2c). Furthermore, the calves had multiple large abscesses in the cranial pulmonary lobes (Fig. 3a). They were accompanied by dystelectasis and a mild suppurative bronchopneumonia in the surrounding lung tissue (Fig. 3). There were subcutaneous masses in the region of the left mandibular lymph node and in front of the shoulder joint (Fig. 1a), and were identified as abscesses (Fig. 4).

Genetic findings

On the farm, the sire of the three affected calves was used in a single breeding period for natural mating with a total of 30 heifers, all originating from a common natural service sire used at the farm in the years before (Fig. 5a). Besides the two affected female calves investigated in this study, 11 more female and 17 male calves were born. One of these additional female calves showed similar signs as observed in the two presented cases and died of pneumonia. However, in this additional case no close clinical examination was done and no sample is available. Neither the 17 males nor the other 10 female offspring showed comparable skin alterations although it was not possible to follow-up the health status of the male calves after they left the farm with 2-weeks of age for fattening. Both, the sire of the affected calves and all 30 dams were apparently normal. The pedigree of the affected calves is in accordance with recessive inheritance. Genetic testing allowed to rule out bovine leukocyte adhesion deficiency (BLAD) as possible cause for the observed symptoms as both cases were tested negative by a commercial laboratory (IFN Schönow GMbH, Bernau).

Adopting the assumption of a possible recessive mode of inheritance, whole-genome sequencing (WGS) followed by single-nucleotide variant (SNV) and small insertion and deletion (InDel) calling resulted in 23 private protein-changing variants (Additional file 3) beside 4754 private non-coding variants (Additional file 4). Due to the strong effect of the putative genetic defect, we hypothesized that a loss-of-function variant affecting the coding sequence of a gene most likely would be responsible for the disorder. They were found to be homozygous exclusively in the two affected animals and absent or occurring heterozygous in 396 control genomes that were sequenced in the course of other ongoing projects at the Institute of Genetics. In a second step 22 variants out of these 23 protein-changing variants were found

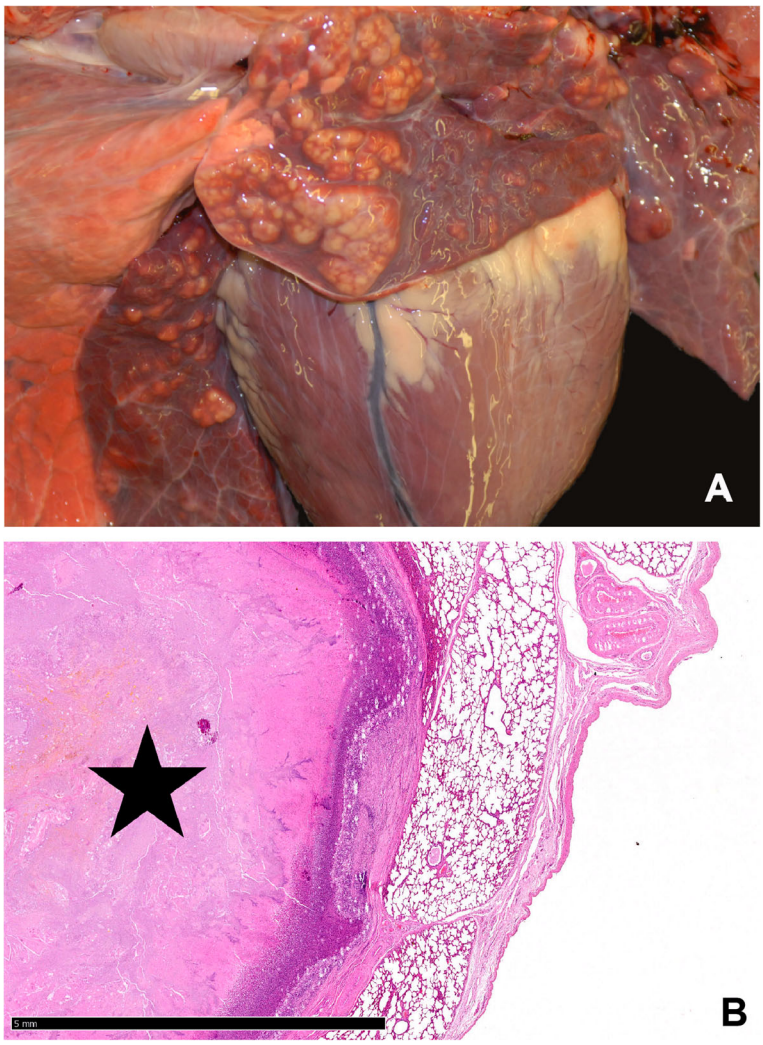


Fig. 3 Lung of case 1. **a** Macroscopic image of multiple abscesses in the cranial pulmonary lobes. **b** Microscopic view of the abscess of the cranial pulmonary lobes (star). Hematoxylin and Eosin, 1.25x, bar = 5 mm

with both possible genotypes (heterozygous and homozygous) in the 3103 control genomes from the 1000 Bull Genomes project (Additional file 3) [17]. The single remaining protein-changing variant in the *interleukin 17 receptor A* (*IL17RA*) gene, which was absent from all controls was predicted to have a deleterious impact on the encoded protein. This single nucleotide deletion in exon 3 of *IL17RA* (chr5: g.108813251delC; Fig. 5b) leads to a frameshift (XM_015460734.2: c.180delC), which is predicted to replace 61 amino acids and to truncate a significant part (~ 85%) of the C-terminus of the protein (XP_015316220.2: p.(Cys61AlafsTer62)).

We genotyped the single-nucleotide deletion in exon 3 of *IL17RA* in both available cases, as well as in eleven non-affected relatives (Fig. 5a) by Sanger sequencing. This confirmed the homozygous state of the variant in the two affected calves, whereas both dams of the affected calves were heterozygous for the *IL17RA* deletion, as well as two half-sibs and two maternal sisters (Fig. 5a). All further tested relatives were homozygous wild type.

Discussion

Interestingly our study revealed *interleukin 17 receptor A* (*IL17RA*) as a new gene involved in a recessively inherited genodermatosis of Holstein cattle. The herein described clinical findings resemble strongly a disorder known in Fleckvieh cattle as zinc deficiency-like (ZDL) syndrome associated with a *PLD4* missense variant [15]. Both herein described cases and the ZDL-affected calves show an increasing distribution and severity of skin lesions over time, where the first lesions appeared on the muzzle. Further common characteristics are the occurrence of pneumonia and the inability to response to oral

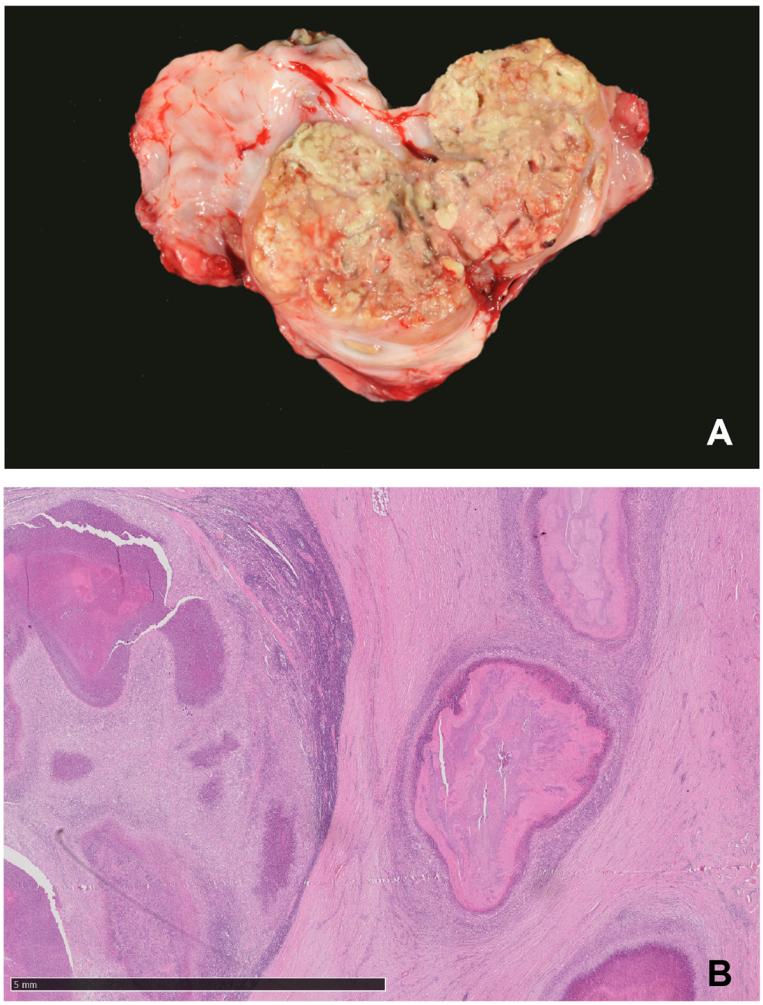


Fig. 4 Submandibular abscess observed in case 1. **a** Macroscopic image of abscess formation at the left mandibular area, orange-sized abscess in the subcutis. **b** Microscopic view showing surrounding granulation tissue, Hematoxlyin and Eosin 1,25x, bar =5 mm

zinc supplementation [16]. However, the *PLD4* variant and the previously reported *SLC39A4*-associated zinc deficiency in Holstein cattle (BHZD) could be ruled out due to their clinical differences. While the dermatological changes associated with the *PLD4* non-sense variant are described as a crusting dermatitis [16], the herein described cases show open wounds that scarcely heal. BHZD affected calves show a zinc deficiency, which can be treated by oral application of zinc [13]. This therapy was applied to the herein presented cases but did not lead to any clinical improvement and thereby a disorder like BHZD could be ruled out. Furthermore, bovine leukocyte adhesion deficiency (BLAD) was excluded through genetic testing. Finally, all known disease-causing variants were confirmed to not be present in the described cases by analysing the WGS data.

To our knowledge, the *IL17RA* gene had not yet been associated with similar disorders in domestic animals. Variants in human *IL17RA* were reported to cause autosomal recessively inherited immunodeficiency 51 (OMIM 605461) and affected individuals show inborn susceptibility to several infections and especially to chronic mucocutaneous candidiasis [18, 19]. The pathogenic variants associated with immunodeficiency 51 include missense variants as well as non-sense variants and lead to a total loss-of-function of IL17RA [19, 20]. Pro-inflammatory cytokines such as interleukin 17 and 22, as well as their receptors play a central role in acute and chronic inflammatory responses [21, 22]. Interleukin 17 (IL-17) is especially known to be of importance at mucosal and barrier surfaces and plays a possible role in the pathogenesis of autoimmunity [23]. The most common member of the IL-17 family is IL-17A and the receptor complex it signalling threw the heterodimer complex formed by the receptors IL-17RA and IL-17RC [23–25]. Furthermore, IL-17 had been associated with a

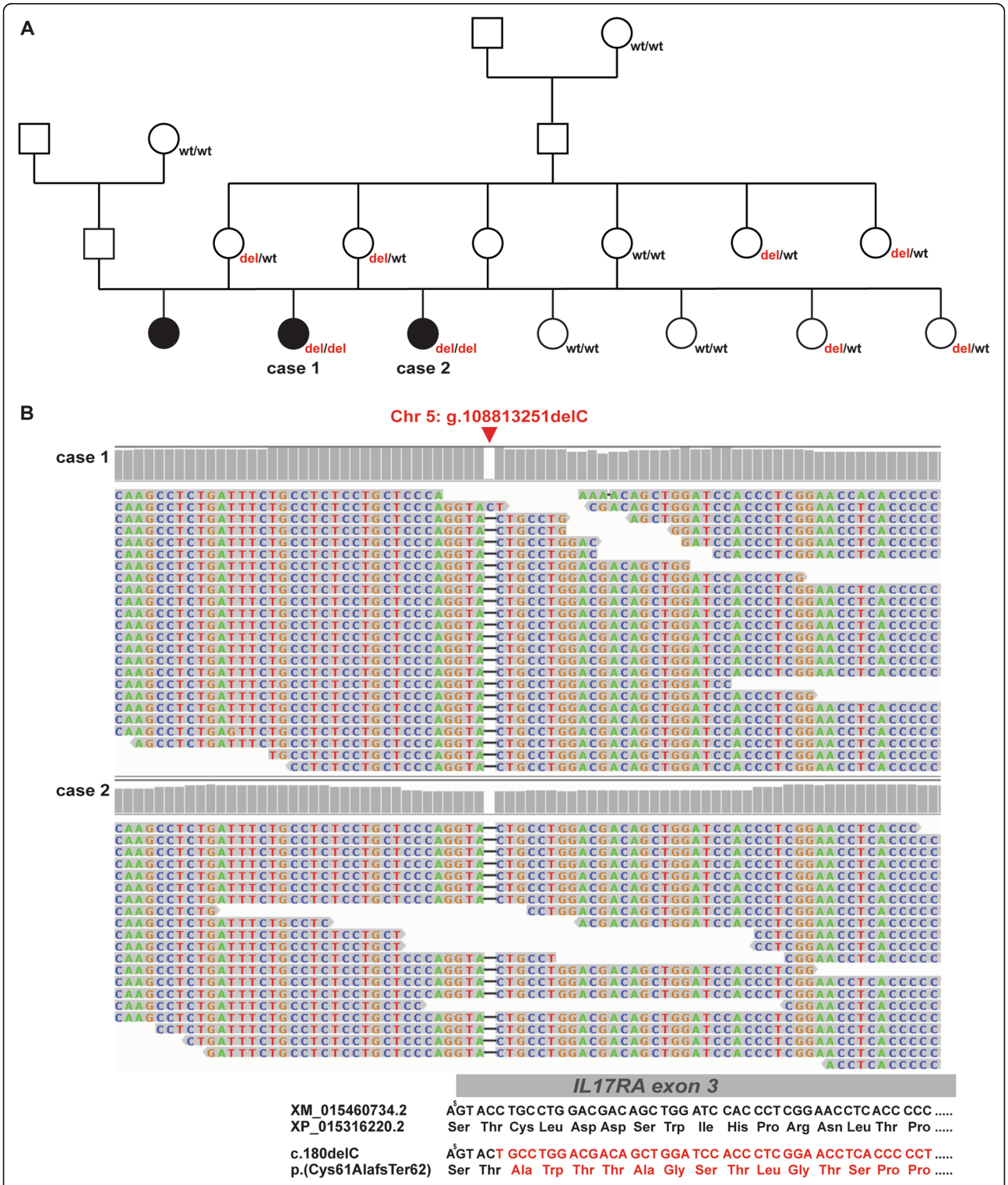


Fig. 5 *IL17RA* frameshift variant associated with the observed genodermatosis. **a** Pedigree including the three affected calves (shown in black). All thirteen available family members were genotyped for the *IL17RA* variant (genotypes presented below the individuals). **b** IGV screenshot showing the single-nucleotide deletion on chromosome 5 in the two sequenced cases. Note the indicated change in the deduced amino acid sequence below. A premature stop codon is predicted to occur after 62 altered residues. ($ indicates the first nucleotide of codon 59 belonging to *IL17RA* exon 2)

wide range of diseases including inflammation ranging from arthritis, psoriasis, spondylitis, Crohn's disease, multiple sclerosis, cardio vascular diseases and a variety of disorders of the lung [23]. The involvement of IL-17 and its receptor IL17RA has been shown in respect to psoriasis and psoriatic arthritis in humans and mice [26–30].

Interestingly, the herein described bovine cases homozygous for the *IL17RA* loss-of-function variant did not suffer from fungal but from multi-bacterial infections. This points to differences in the immune defence against fungal and bacterial infections between humans/mice and cattle.

This *IL17RA*-associated semi lethal genodermatosis, leading to psoriasis-like skin alterations and immunodeficiency resembling zinc deficiency-like syndrome in Holstein cattle, was not reported before. Based on the absence of the disease-causing *IL17RA* variant in the global cohort of whole-genome sequenced cattle, which includes several hundred Holstein sires used in artificial insemination, the frequency of the mutant allele in the international Holstein population is probably very low. Nonetheless, our findings allow for the first time a targeted monitoring of the prevalence of this most likely pathogenic variant in the local German population and the avoidance of further risk matings.

Conclusion

To the best of our knowledge, a similar *IL17RA*-associated genodermatosis showing chronic skin alterations correlated with an inherited immunodeficiency has not been described until now in cattle. Moreover, this is the first report of a most likely pathogenic variant in *IL17RA* in a domestic animal species.

Methods

Animals

In a single German Holstein breeding flock three calves which were sired by the same bull showed clinical signs suspicious for a disease similar to ZDL or BHZD. While one calf perished of a heavy lung inflammation during the first weeks of life the other two calves were chronically suffering from psoriasis-like skin alterations and showed reduced growth. Therefore, in this study clinical and pathological examinations were conducted on the two affected calves and the genetic background. Blood samples from eleven healthy relatives were collected. Unfortunately, the sire was already slaughtered when the study was executed. In addition, we collected blood of two dams, the paternal grandmother, the maternal great-grandmother, four paternal half-sibs and three paternal half-sibs of the dams (Fig. 5a).

Clinical analyses

Clinical examination was conducted on two affected calves. This includes a general description of their health status. Further, all important peripheral blood parameters were analysed. Blood analyses were performed using EDTA, serum and lithium-heparin samples from each of the affected calves. After necropsy microbiological analysis of the skin and lung tissue had been performed.

Whole-genome sequencing

In order to investigate the genetic architecture of the disorder, genomic DNA was extracted from EDTA blood samples of case 1 and 2 and used for whole-genome sequencing (WGS). Therefore, individual PCR-free fragment libraries, which were sequenced for 150 bp paired-end reads were prepared. Both cases were sequenced on the Illumina NovaSeq6000 resulting in a read depth of approximately 17.2x and 20.8x for the first and the second case, respectively. The WGS data was mapped to the latest reference genome ARS-UCD1.2 [31] and called for small nucleotide variants (SNVs) and small insertions and deletions (InDels). We followed the workflow proposed by the 1000 Bull Genomes Project (run 7) [17, 32] to process the raw data into binary alignment map (BAM) and genomic variant call format (GVCF) files. Further, the individual GVCF files were merged to one large variant call format (VCF) file by using CombineGVCFs and CatVariants of GATK v3.8 [33]. This file was produced together with another 396 genomes available at the Institute of Genetics of the University of Bern and thereby includes 398 genomes (Additional file 5) of a variety of 23 cattle breeds and few crossbred animals. For this VCF file SNVs and InDels were called using GenotypeGVCF of GATK v3.8 [33]. Furthermore, quality labels based on the best practice recommendations in GATK using the VariantFiltration of GATK v3.8 [33] were given for each variant. Finally, SnpEff v4.3 [34] was used to functionally annotated each variants effect, based on the NCBI Annotation Release 106 [35]. This resulted in the final VCF file, which includes all 398 individuals' variants and their functional annotation.

We filtered the final VCF file based on the assumption that the disorder is a rare recessive disorder. Thereby, only the two affected animals need to be carriers of a possible candidate variant but no other of the 396 control genomes (Additional file 5). This approach is supported by the fact, that there are no other sequenced German Holstein animals in the data set. In order to validate the allele frequency of the variant and exclude variants that occur in other breeds all remaining variants were compared to the VCF file from the 1000 Bull Genomes project (run 7) [17]. This reference data includes 3103 cattle genomes which include 937 individuals from the Holstein breed from all around the globe. Integrative

Genomic Viewer (IGV) software [36] was used for visual inspection of the remaining candidate variants.

Genotyping of the candidate variants

Forward primer 5′-GTCATGGCCTGACTGTGAAG-3′ and reverse primer 5′-GTCCACTCGATGTGAACCAC-3′ were designed with the software Primer3 [37] to produce a 243 bp fragment including the *IL17RA* variant. Sanger sequencing of the resulting PCR product was performed by a service laboratory (LGC Genomics, Berlin). The obtained sequences were analysed with ChromasPro 1.22 software (Technelysium Pty Ltd., South Brisbane) and compared with the relevant genome region of the latest reference genome ARS-UCD1.2 [31].

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12863-020-00860-4.

Additional file 1. Results of peripheral blood analyses.

Additional file 2. Results of microbiological analyses in lung and skin samples. Note that the analyses could not display any mycological agents.

Additional file 3. Twenty-three shared homozygous protein-changing variants in both sequenced cases. The frequency of the corresponding genotypes in the variant catalogue of the 1000 Bulls Genome project are given.

Additional file 4. 4754 shared homozygous non-coding variants in both sequenced cases.

Additional file 5. EBI accession numbers of all publicly available genome sequences.

Abbreviations

BAM: Binary alignment map; BHZD: Bovine hereditary zinc deficiency; BLAD: Bovine leukocyte adhesion deficiency; BPV: Bovine poxvirus; BTV: Bluetongue virus; BVD: Bovine virus diarrhoea; GVCF: Genomic variant call format; InDel: Small insertion and deletion; LSD: Lumpy skin disease; SNV: Single-nucleotide variant; VCF: Variant call format; WGS: Whole-genome sequencing; ZDL: Zinc deficiency-like

Acknowledgements

The authors are grateful to the cattle breeder for submission of the affected animals. We acknowledge the assistance of the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome re-sequencing experiments. Furthermore, we thank the Interfaculty Bioinformatics Unit of the University of Bern for providing high performance computing infrastructure.

Authors' contributions

MS conducted the clinical examinations and clinical laboratory analysis. IH performed the bioinformatics of the whole-genome sequencing data and drafted the manuscript. MHO ruled out infectious and heritable differential diagnosis, collected samples on farm, advised the farmer to submit the animals to the veterinary clinic of the Justus-Liebig University and contributed pedigree information. MHE and LMR performed the pathological examinations. SM, the veterinarian attending the farm, examined and treated the affected calves on farm and was the first to suspect a potential inherited disease. CD planned and supervised the whole-genome sequencing analysis. GL conducted pedigree analysis and genotyping of potential causal variants in relatives of the affected calves and in unrelated animals. MS, GL and CD coordinated the study and contributed to the writing of the manuscript. All authors have read and approved the final manuscript.

Funding

This study was partially funded by the Swiss National Science Foundation.

Availability of data and materials

Whole-genome sequencing data can be accessed on the European Nucleotide Archive with the project ID PRJEB18113 (sample accessions: SAMEA5714977; SAMEA5714978). Further publicly available data used in this study can be found in this repository with the indicated sample IDs in the additional file 4.

Ethics approval and consent to participate

This study did not require official or institutional ethical approval as it was not experimental but part of clinical and pathological veterinary diagnostics. The animals were handled according to good ethical standards and German legislation (German Animal Welfare Act from 15th May 2006, last amended on 20th November 2019). A formal ethics approval for the euthanasia of the affected calves was not needed as this complies with §4 of the German Animal Welfare Act. The affected calves were euthanized by intravenous injection of an overdose of sodium pentobarbital (500 mg/mL), performed by MS. The aim was to identify the cause of the congenital disorder and thereby improve the animal welfare situation on the farm. Written consent was obtained from the cattle owner.

Consent for publication

Not applicable.

Competing interests

The authors declare that they no competing interests.

Author details

[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. [2]Clinic for Obstetrics, Gynecology and Andrology of Large and Small Animals with Ambulatory Service, Faculty of Veterinary Medicine, Justus-Liebig University Giessen, 35392 Giessen, Germany. [3]Bovine Health Service, Chamber of Agriculture of North Rhine-Westphalia, 59505 Bad Sassendorf, Germany. [4]Institute of Veterinary Pathology, Faculty of Veterinary Medicine, Justus-Liebig-University Giessen, 35392 Giessen, Germany. [5]Veterinary Sharing Practice, Dr. Siegfried Marquardt and Peter Walter, 47574 Goch, Germany. [6]Institute of Animal Breeding and Genetics, Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Justus-Liebig University Giessen, 35390 Giessen, Germany.

Received: 27 February 2020 Accepted: 14 May 2020
Published online: 24 May 2020

References

1. Alsaaod M, Locher I, Jores J, Grimm P, Brodard I, Steiner A, et al. Detection of specific Treponema species and Dichelobacter nodosus from digital dermatitis (Mortellaro's disease) lesions in Swiss cattle. Schweiz Arch Tierheilkd. 2019;161(4):207–15.
2. Bianchi MV, Silveira S, Mosena ACS, de Souza SO, Konradt G, Canal CW, et al. Pathological and virological features of skin lesions caused by BVDV in cattle. Braz J Microbiol. 2019;50(1):271–7.
3. Beard PM. Lumpy skin disease: a direct threat to Europe. Vet Rec. 2016; 178(22):557–8.
4. Marin RE, Houston R, Omanska-Klusek A, Alcaraz A, Garcia JP, Uzal FA. Pathology and diagnosis of proliferative and ulcerative dermatitis associated with Tunga penetrans infestation in cattle. J Vet Diagn Investig. 2015;27(1): 80–5.
5. Coppock RW, Mostrom MS, Simon J, McKenna DJ, Jacobsen B, Szlachta HL. Cutaneous ergotism in a herd of dairy calves. J Am Vet Med Assoc. 1989; 194(4):549–51.
6. Kielland C, Ruud LE, Zanella AJ, Osteras O. Prevalence and risk factors for skin lesions on legs of dairy cattle housed in freestalls in Norway. J Dairy Sci. 2009;92(11):5487–96.
7. Chermette R, Ferreiro L, Guillot J. Dermatophytoses in animals. Mycopathologia. 2008;166(5–6):385–405.
8. Medeiros GX, Riet-Correa F. Epidermolysis bullosa in animals: a review. Vet Dermatol. 2015;26(1):3–13 e1-2.
9. Hensel P. Nutrition and skin diseases in veterinary medicine. Clin Dermatol. 2010;28(6):686–93.

10. Leeb T, Müller EJ, Roosje P, Welle M. Genetic testing in veterinary dermatology. Vet Dermatol. 2017;28(1):4–e1.
11. Shuster DE, Kehrli ME, Ackermann MR, Gilbert RO. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. Proc Natl Acad Sci. 1992 Oct 1;89(19):9225–9.
12. Nagahata H. Bovine leukocyte adhesion deficiency (BLAD): a review. J Vet Med Sci. 2004;66(12):1475–82.
13. Machen M, Montgomery T, Holland R, Braselton E, Dunstan R, Brewer G, et al. Bovine hereditary zinc deficiency: lethal trait a 46. J Vet Diagn Investig. 1996;8(2):219–27.
14. Yuzbasiyan-Gurkan V, Bartlett E. Identification of a unique splice site variant in SLC39A4 in bovine hereditary zinc deficiency, lethal trait A46: an animal model of acrodermatitis enteropathica. Genomics. 2006;88(4):521–6.
15. Jung S, Pausch H, Langenmayer MC, Schwarzenbacher H, Majzoub-Altweck M, Gollnick NS, et al. A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. BMC Genomics. 2014;15:623.
16. Langenmayer MC, Jung S, Majzoub-Altweck M, Trefz FM, Seifert C, Knubben-Schweizer G, et al. Zinc deficiency-like syndrome in Fleckvieh calves: clinical and pathological findings and differentiation from bovine hereditary zinc deficiency. J Vet Intern Med. 2018;32(2):853–9.
17. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. Annu Rev Anim Biosci. 2019;7:89–102.
18. Ling Y, Puel A. IL-17 and infections. Actas Dermosifiliogr. 2014;105:34–40.
19. Lévy R, Okada S, Béziat V, Moriya K, Liu C, Chai LY, et al. Genetic, immunological, and clinical features of patients with bacterial and fungal infections due to inherited IL-17RA deficiency. Proc Natl Acad Sci. 2016; 113(51):E8277–E85.
20. Puel A, Cypowyj S, Bustamante J, Wright JF, Liu L, Lim HK, et al. Chronic mucocutaneous candidiasis in humans with inborn errors of interleukin-17 immunity. Science. 2011;332(6025):65–8.
21. de Masson A, Bouaziz JD, Battistella M, Bagot M, Bensussan A. Immunopathology of psoriasis: from bench to bedside. Med Sci. 2016;32(3): 253–9.
22. Gu C, Wu L, Li X. IL-17 family: cytokines, receptors and signaling. Cytokine. 2013;64(2):477–85.
23. Miossec P, Kolls JK. Targeting IL-17 and TH17 cells in chronic inflammation. Nat Rev Drug Discov. 2012;11(10):763–76.
24. Abusleme L, Moutsopoulos NM. IL-17: overview and role in oral immunity and microbiome. Oral Dis. 2017;23(7):854–65.
25. Gaffen SL. Structure and signalling in the IL-17 receptor family. Nat Rev Immunol. 2009;9(8):556–67.
26. Adamopoulos IE, Suzuki E, Chao CC, Gorman D, Adda S, Maverakis E, et al. IL-17A gene transfer induces bone loss and epidermal hyperplasia associated with psoriatic arthritis. Ann Rheum Dis. 2015;74(6):1284–92.
27. Batalla A, Coto E, Gonzalez-Lara L, Gonzalez-Fernandez D, Gomez J, Aranguren TF, et al. Association between single nucleotide polymorphisms IL17RA rs4819554 and IL17E rs79877597 and psoriasis in a Spanish cohort. J Dermatol Sci. 2015;80(2):111–5.
28. Floudas A, Saunders SP, Moran T, Schwartz C, Hams E, Fitzgerald DC, et al. IL-17 receptor a maintains and protects the skin barrier to prevent allergic skin inflammation. J Immunol. 2017;199(2):707–17.
29. Li Y, Golden JB, Camhi MI, Zhang X, Fritz Y, Diaconu D, et al. Protection from psoriasis-related thrombosis after inhibition of IL-23 or IL-17A. J Invest Dermatol. 2018;138(2):310–5.
30. Su Y, Huang J, Zhao X, Lu H, Wang W, Yang XO, et al. Interleukin-17 receptor D constitutes an alternative receptor for interleukin-17A important in psoriasis-like skin inflammation. Sci Immunol. 2019;4(36).
31. NCBI Assembly - ARS-UCD1.2 reference sequence. National Center for Biotechnology Information, U.S. National Library of Medicine. 2018. https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1. Accessed 21 Aug 2018.
32. 1000 Bull Genomes Project. http://www.1000bullgenomes.com/. (2019) Accessed 27 Sept 2019.
33. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491–8.
34. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. FlY. 2012;6(2):80–92.
35. NCBI Genome - Bos taurus Annotation Release 106. National Center for Biotechnology Information, U.S. National Library of Medicine. 2018. https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/. Accessed 21 Aug 2018.
36. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–6.
37. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115.

Publisher's Note

DSP missense variant in a Scottish Highland calf with congenital ichthyosis, alopecia, acantholysis of the tongue and corneal defects

Journal:	BMC Veterinary research
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, genetic figures, co-writing
Displayed version:	published version
DOI:	10.1186/s12917-021-03113-3

BMC Veterinary Research

RESEARCH

Open Access

DSP missense variant in a Scottish Highland calf with congenital ichthyosis, alopecia, acantholysis of the tongue and corneal defects

Irene M. Häfliger[1†], Caroline T. Koch[1†], Astrid Michel[2], Silvia Rüfenacht[3,4], Mireille Meylan[2], Monika M. Welle[5] and Cord Drögemüller[1*]

Abstract

Background: Ichthyosis describes a localized or generalized hereditary cornification disorder caused by an impaired terminal keratinocyte differentiation resulting in excessive stratum corneum with the formation of more or less adherent scales. Ichthyosis affects humans and animals. Two rare bovine forms are reported, the severe harlequin ichthyosis and the less severe congenital ichthyosis, both characterized by a severe orthokeratotic lamellar hyperkeratosis.

Results: A 2-weeks-old purebred Scottish Highland calf was referred because of a syndrome resembling congenital ichthyosis. The clinical phenotype included diffuse alopecia and a markedly lichenified skin covered with large and excessive scales. Additionally, conjunctivitis and ulceration of the cornea were noted. Post-mortem examination revealed deep fissures in the diffusely thickened tongue and histopathological findings in the skin confirmed the clinical diagnosis. Whole-genome sequencing of the affected calf and comparison of the data with control genomes was performed. A search for private variants in known candidate genes for skin phenotypes including genes related with erosive and hyperkeratotic lesions revealed a single homozygous protein-changing variant, *DSP*: c.6893 C>A, or p.Ala2298Asp. The variant is predicted to change a highly conserved residue in the C-terminal plakin domain of the desmoplakin protein, which represents a main intracellular component of desmosomes, important intercellular adhesion molecules in various tissues including epidermis. Sanger sequencing confirmed the variant was homozygous in the affected calf and heterozygous in both parents. Further genotyping of 257 Scottish Highland animals from Switzerland revealed an estimated allele frequency of 1.2%. The mutant allele was absent in more than 4800 controls from various other cattle breeds.

Conclusions: This study represents the first report of combined lesions compatible with congenital ichthyosis, alopecia, acantholysis of the tongue and corneal defects associated with a *DSP* missense variant as the most likely underlying cause. To the best of our knowledge, this study is also the first report of a *DSP*-related syndromic form of congenital ichthyosis in domestic animals. The results of our study enable genetic testing to avoid the unintentional occurrence of further affected cattle. The findings were added to the Online Mendelian Inheritance in Animals (OMIA) database (OMIA 002243-9913).

*Correspondence: cord.droegemueller@vetsuisse.unibe.ch
†Irene M. Häfliger and Caroline T. Koch contributed equally to this work.
[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland
Full list of author information is available at the end of the article

Keywords: Cattle, Dermatology, Skin, Genodermatosis, Corneal ulcers, Hyperkeratosis, Genetic disorder, Precision medicine

Background

Ichthyoses encompass a heterogeneous group of congenital disorders characterized by an abnormal terminal keratinocyte differentiation [1]. They are linked by the common finding of a thickened stratum corneum resulting in localized or generalized scaling. In ichthyosis, desquamation of corneocytes is impaired, resulting in retained large squames and thus severe hyperkeratosis, as well as loss of skin elasticity and an abnormal barrier function [1, 2]. The thick skin with an excessive amount of superficial scales resembles fish scales (greek ichthys means fish) and gave the name for these disorders. In 2009, the various ichthyotic diseases in people were classified in a consensus paper [2]. In that milestone article, the classification is based on clinical findings and the inherited ichthyoses are subdivided into the large groups of syndromic versus non-syndromic forms. The further subdivision distinguishes between the common ichthyosis (ichthyosis vulgaris), autosomal recessive congenital ichthyosis, and keratinopathic ichthyosis [3]. The large group of recessively inherited disorders with congenitally appearing ichthyosis but no extra-cutaneous involvement is heterogeneous and can be subdivided in three clinical phenotypes: (1) Harlequin ichthyosis representing the most severe, mostly lethal phenotype; (2) congenital ichthyosis associated with erythema and fine white scales; and (3) lamellar ichthyosis characterized by large dark scales [3]. In humans, the various forms of ichthyosis are associated with variants in at least 50 genes encoding structural proteins and enzymes affecting several cellular functions including DNA repair, lipid biosynthesis, adhesion, desquamation, as well as other pathways [1–4].

In animals, several forms of non-syndromic presentations of ichthyosis have been reported and clinical signs are present at birth or rarely later in life [5]. A classification system similar to humans does not exist in animals, but localized and generalized forms are known and, based on histology, non-epidermolytic forms are distinguished from epidermolytic forms. Various forms of non-syndromic, mostly inherited forms of ichthyosis have been identified in domestic animals [5–12]. In cattle, forms of ichthyosis fetalis, which represents a severe form also described as harlequin-like ichthyosis (OMIA 002238-9913) and congenital ichthyosis (OMIA 001993-9913) have been reported. In older reports, the diagnosis was based only on the clinical and/or histopathological findings. In cattle, affected animals show variable degrees of generalized hyperkeratosis; large thick cutaneous scales situated mainly to the inguinal region, limbs, abdomen and muzzle are typical [8, 9]. In some syndromic cases, microtia, cataracts and thyroid abnormalities have been reported as well [12]. Calves affected with congenital ichthyosis show milder but comparable lesions to those of ichthyosis fetalis [10]. Changes in the hair coat such as hypotrichosis or alopecia are clinically described but may represent secondary findings.

The underlying genetic causes for ichthyosis fetalis are mostly unknown in cattle, except for an *ABCA12*-related harlequin-like form (; OMIA 002238-9913). This recessively inherited lethal disorder observed in Chianina cattle is characterized by a hyperkeratotic and significantly lichenified skin with widely distributed, erytematous, deep fissures [7, 8, 13]. Two further pathogenic loss-of-function variants in the bovine *ABCA12* gene associated with nonviable forms of ichthyosis fetalis have been reported in Shorthorn [14] and Hereford cattle [15]. In addition, a deleterious frameshift variant in the *FA2H* gene of Chianina cattle associated with a mild syndromic form of bovine ichthyosis congenita (OMIA 002450-9913) has recently been reported [16].

At present, this obvious genetic heterogeneity of inherited cornification disorders can be analyzed in cattle using whole-genome sequencing (WGS)-based precision diagnostics [17]. Studying the molecular aetiology of single cases is nowadays also feasible in cattle [18–20]. Therefore, the purpose of this study was to characterize the clinical and pathological phenotype of an ichthyosis-affected Scottish Highland calf, and to evaluate its possible genetic etiology using WGS.

Methods

WGS was performed on DNA extracted from ethylenediaminetetraacetic acid (EDTA) blood of the calf. An individual PCR-free fragment library with approximately 400 bp inserts was created and sequenced on a NovaSeq6000 for 150 bp paired-end reads (Illumina, San Diego, CA, USA). The sequenced reads were mapped to the ARS-UCD1.2 reference genome [21] resulting in an average read depth of approximately 19-fold, and single-nucleotide variants and small indel variants were called in accordance as described before [22]. The applied software and steps to process fastq files into binary alignment map and genomic variant call format files were in accordance with the processing guidelines of the 1000 Bull Genomes Project [23]. The effects of the called variants were functionally evaluated with snpeff v4.3 [24],

using the NCBI annotation release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/). In order to find private variants, we compared the genotypes of the affected calf with 705 cattle genomes of various breeds that had been sequenced in the course of the Swiss Comparative Bovine Resequencing project [25]. An *in silico* assessment of the molecular consequences of the identified amino acid exchanges was carried out with PROVEAN [26].

Results

Clinicopathological findings

A 2-week-old female Highland cattle calf weighing 51 kg was presented to the Clinic for Ruminants of the Vetsuisse Faculty of the University of Bern with severe skin lesions (Fig. 1). The calf presented with diffuse alopecia involving approximately 90% of the body. The remaining hair was found on the head, the limbs and the tail, and was easy to pull. The markedly lichenified

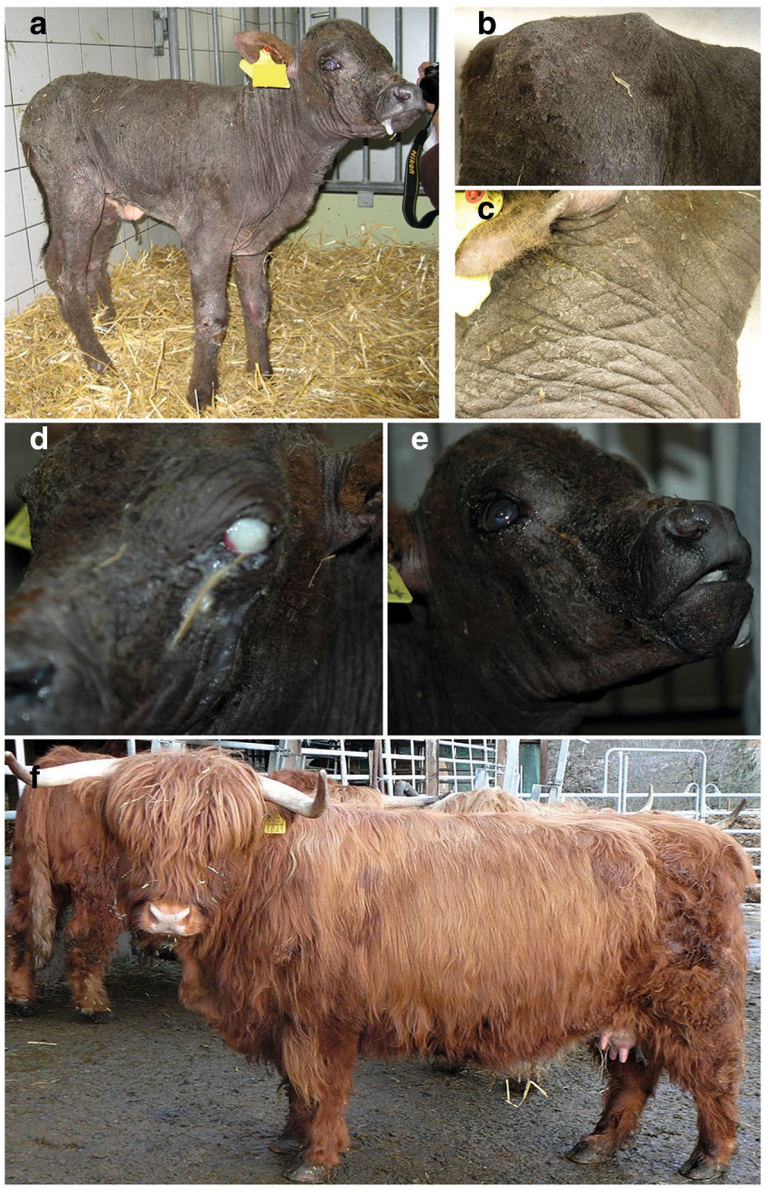


Fig. 1 Clinical phenotype of a 2-week-old Scottish Highland calf with congenital ichthyosis, alopecia and corneal defects. Note the flat crusts lateral to the tarsi, dorsal to the metacarpi and on the bride of the nose (**a**). Backside of the calf with alopecia and large scales (**b**). Alopecia and lichenification on the neck (**c**). Cloudy left eye with conjunctivitis (**d**). The right eye shows a corneal ulceration with a prolapsed iris (**e**). Appearance of a typical purebred Scottish Highland cow (**f**)

hyperkeratotic skin was covered by a thick layer of keratin, which exfoliated as large scales (Fig. 1c). The skin presented multiple folds (Fig. 1a). In addition, the areas lateral to the tarsi, dorsal to the metacarpi and the bridge of the nose were covered with flat crusts (Fig. 1a). The eyelids showed hyperkeratosis as well (Fig. 1d). The ears were of normal length and the mucocutaneous junctions were unaltered (Fig. 1a). The epithelium on the dorsal surface of the tongue was diffusely thickened and presented with multifocal fissures (Additional file 1: Fig. S1). Necrotic debris was covering the surface of the tongue multifocally. There were whitish plaque-like deposits on the ventral tip of the tongue (Fig. 1d). The hoofs seemed intact.

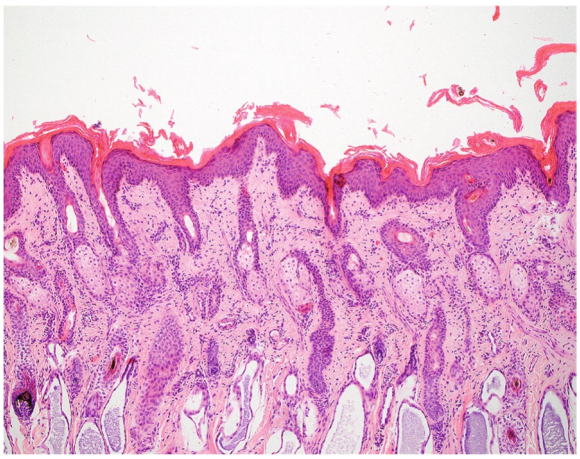


Fig. 2 Histological investigation of the skin of a 2-week-old Scottish Highland calf with congenital ichthyosis and alopecia. Note the compact to laminated orthokeratotic hyperkeratosis extending into the follicular infundibuli and the mild to moderate hyperplasia of the epidermis. Haematoxylin and eosin stain, x100

Due to the severity of the skin lesions and the poor prognosis, the calf was euthanized one day after admission and necropsy with subsequent histological examination of various tissues including the skin and the tongue was performed. Histologically, the epidermis was mildly to moderately hyperplastic and covered with abundant compact to laminated keratin and multifocal crusts (Fig. 2). The hyperkeratosis extended into the follicular infundibulum and sometimes into the ducts of the sebaceous glands. Many hair follicles were dysplastic, presenting with a false orientation of the infundibula and isthmic part, a distorsion of the inferior portion, a multifocally thickened outer root sheath and hair shafts, which were either broken or had an irregular contour. In addition, there was a mild perivascular infiltrate composed mostly of lymphocytes and plasma cells in the superficial dermis.

The epithelium of the tongue was diffusely hyperplastic and presented with multifocal deep fissures (Fig. 3a). The keratinocytes in the stratum spinosum of the tongue epithelium were rounded and towards the surface adhesion to the neighboring keratinocytes was completely lost (acantholysis). Necrotic debris was present on the surface. On the tongue, the mucosa of the dorsal surface showed a marked parakeratotic hyperkeratosis (Fig. 3a). A mild infiltration of lymphocytes and plasma cells in the lamina propria was seen as well (Fig. 3b). The fissures presented histologically as full thickness ulceration and serocellular crust formation.

The cornea of the right eye presented with a focal perforating ulcer of 5 mm in diameter with a partial prolapse of the iris through the ulcerated surface and attachment of the iris to the cornea (Fig. 1d). Histologically the cornea neighboring the transmural ulcer was infiltrated with lymphocytes and neutrophils. Bacteria were also present. In addition, new vessel formation in the cornea was seen. The iridocorneal angle was constricted and contained

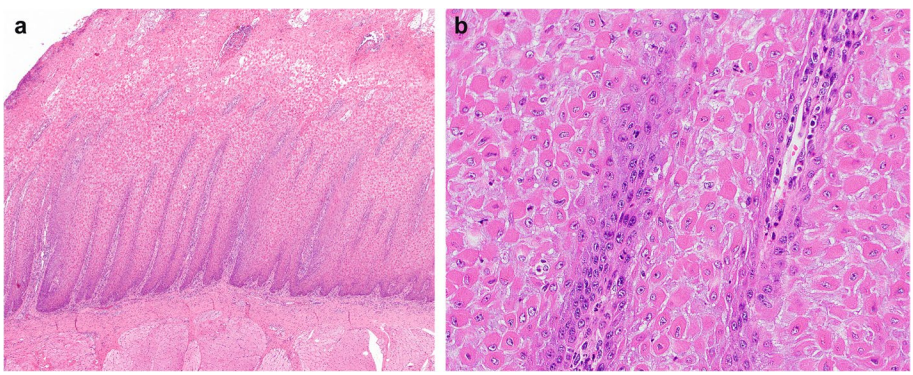


Fig. 3 Histological investigation of the tongue of a two-week-old Scottish Highland calf with congenital acantholysis of the tongue. Note the severely hyperplastic epithelium and the rounded keratinocytes losing the adhesion to the neighbouring cells (**a**) as well as the mild infiltration of lymphocytes and plasma cells in the lamina propria (**b**). Hematoxylin and eosin stain, x100 (**a**), x400 (**b**)

many neutrophils. The retina was detached. The left eye was macroscopically cloudy and conjunctivitis was present (Fig. 1e). Histopathological examination of the left eye revealed a transmural ulceration of the cornea and rupture of the Descemet's membrane. Large amounts of fibrin, neutrophils and bacteria were found in the anterior chamber. The iris was attached to the cornea and the retina was diffusely ablated.

Besides these major pathological findings, a mild acute purulent bronchopneumonia and a diffuse severe subacute pustular rumenitis without hyperkeratosis were noted. Other tissues were unremarkable. In summary, the clinical and pathological findings in the skin of this calf were consistent with congenital ichthyosis and a follicular / hair shaft dysplasia. The findings in the tongue and the eyes are remarkable and have, to the best of our knowledge, not been described in association with similar skin lesions.

Pedigree analysis

The studied calf was the only affected animal in a Swiss herd of purebred Scottish Highland cattle. The sire of the present case was a natural service purebred Scottish Highland bull, which sired further 105 normal offspring within eight years. An enquiry with the Scottish Highland cattle breeders in Switzerland revealed no evidence of other similar cases in the past. The available pedigree records of the calf's ancestry were analyzed and multiple inbreeding loops between the parents were found (Fig. 4). We detected at least five common ancestors occurring 7-8 generations ago. In light of the obvious consanguinity as well as both parents are unaffected, we hypothesized that the current case might be explained by a rare recessively inherited variant. Nonetheless, due to the sporadic occurrence a *de novo* mutation in the germline of one parent or during very early embryonic development could not be fully excluded.

Genetic findings

Based on the assumed recessive inheritance and a sublethal effect of the disease-causing variant, we hypothesized that, most likely, a loss-of-function mutation affecting a protein-coding gene would be responsible for the observed disorder. Therefore, we subsequently concentrated on protein-changing variants with a moderate or high predicted impact on the encoded protein. This revealed five protein-changing variants, located within different genes or loci, exclusively present homozygous in the genome of the ichthyosis-affected calf (Additional file 2: Table S1). Of all these five remaining private variants, only one occurred in a candidate for ichthyosis: desmoplakin (*DSP*). The variant can be designated as Chr23:47826600G>T. It

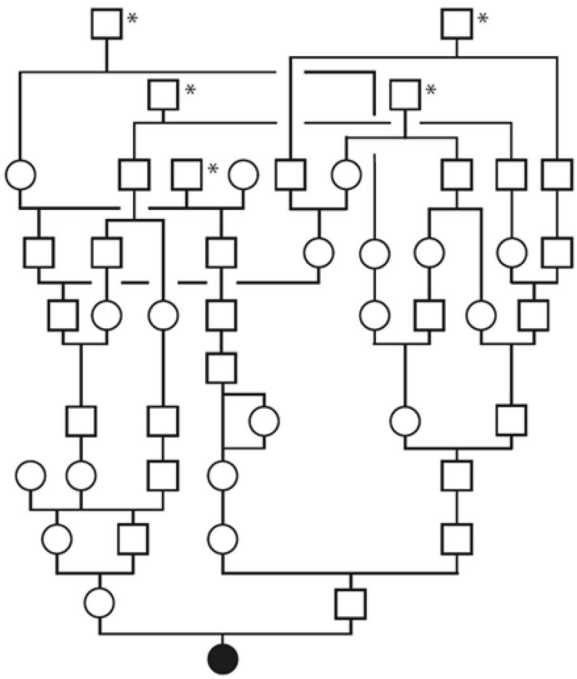


Fig. 4 Pedigree of the Scottish Highland calf with a syndromic form of congenital ichthyosis. Note the multiple inbreeding loops between the parents. Five common ancestors are marked by an asterisk. The affected calf is marked with a black circle. Females are indicated as circles and males as squares

is a missense variant, NM_001192368.2:c.6893 C>A, predicted to change a highly conserved alanine residue at the C-terminal plakin domain affecting the second plakin-repeat subdomain of desmoplakin, NP_001179297.1:p.(Ala2298Asp). The *DSP* missense variant was predicted to be deleterious (Table S1).

This variant affecting a candidate gene for loss of keratinocyte adhesion explains the acantholysis in the tongue of the calf well. It may also affect the corneodesmosomes in the epidermis and is likely the pathogenic variant for the observed phenotype. We confirmed the presence of the *DSP* missense variant by Sanger sequencing (Fig. 5). The mutant *DSP* allele showed the expected co-segregation with ichthyosis in the available family trio, this was not the case for the other four protein-changing variants in which both parents also showed the alternative homozygous genotype (Table S1). Genotyping of 257 Scottish Highland cattle revealed no homozygous mutant animal and a total of six heterozygous carriers including both parents confirming recessive inheritance. Furthermore, the variant was absent in 4109 cattle genomes of a variety of global breeds including 12 Scottish Highland animals that were part of the run 8 of the ongoing 1000 Bull Genomes Project [23].

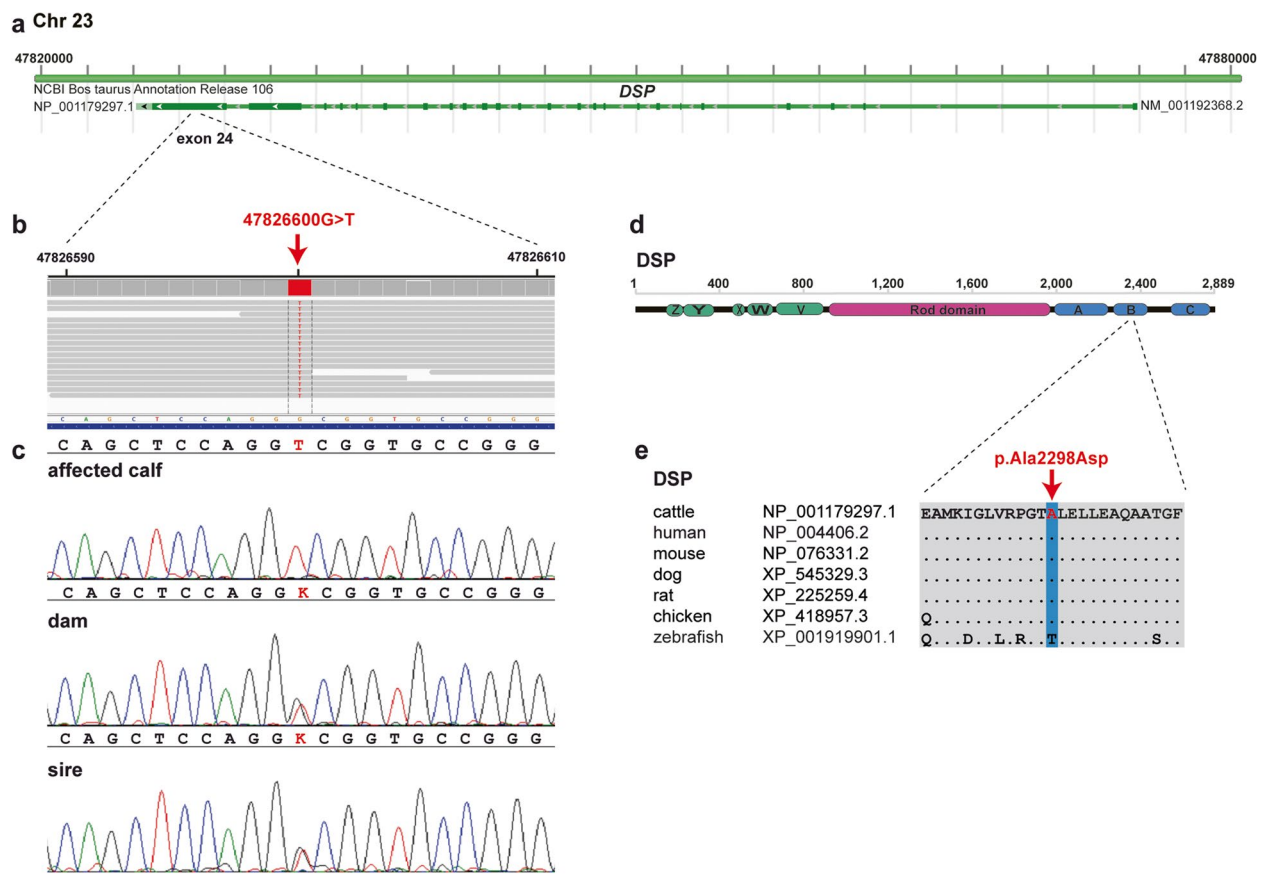


Fig. 5 A missense variant in the *desmoplakin* (*DSP*) gene is associated with a syndromic form of congenital ichthyosis in Scottish Highland cattle. Schematic diagram of the *DSP* gene showing the location of the candidate pathogenic variant in exon 24 (**a**). Screenshot of the WGS data shows the homozygous single nucleotide variant (red arrow) present in the affected calf (**b**). Electropherograms from Sanger sequencing data (**c**). Note that the DNA sample from the affected calf carried the mutant allele in a homozygous state, while the parents were heterozygous, as expected for obligate carriers. The bovine desmoplakin protein is composed of the N-terminal plakin domain with 5 α-helical bundles (Z, Y, X, W, V), the coiled-coiled rod domain, and the C-terminal plakin domain with the 3 plakin-repeat subdomains (A, B, C) (**d**). Multiple-species DSP protein alignment (**e**). The predicted change of the highly conserved alanine to aspartic acid in the affected calf is highlighted by a red arrow

Discussion

The aim of this study was to characterize the clinicopathological phenotype and the genetic aetiology of the observed lesions in the skin, the tongue and the eyes of a Highland calf. The skin lesions are compatible with congenital ichthyosis, which has never been reported in a purebred Scottish Highland calf. Herein we also present evidence for the occurrence of a novel form of recessively inherited ichthyosis due to a homozygous missense variant in the bovine *DSP* gene, which enables selection against this disorder. We hypothesize that this most likely pathogenic deleterious variant is also causative for the alopecia, the erosive tongue lesions due to severe acantholysis and the corneal ulcers. Due to the consanguinity and in light of the syndromic disease phenotype, there is theoretically the possibility of a second genetic disorder; however, the genome sequencing carried out did not reveal any evidence of a second pathogenic protein-changing variant.

Desmoplakin is part of all desmosomes, which are abundantly expressed in both myocardial and epidermal tissue and serve as intercellular adhesion molecules to resist mechanical stress. It anchors intermediate filaments to desmosomal plaques and forms an obligate component of functional desmosomes. Mutations in genes encoding for desmosomal components are associated with a broad spectrum of phenotypes comprising skin and hair abnormalities, and account for 45-50% of cases of arrhythmogenic right ventricular cardiomyopathy in humans [27]. More than 120 dominant and recessive *DSP* variants have been reported to be associated with skin, hair and/or heart defects such as dominant inherited arrhythmogenic right ventricular

cardiomyopathy (OMIM 615,821) and recessive inherited Carvajal syndrome (OMIM 605,676) characterized by an extreme type of dominant arrhythmogenic right ventricular cardiomyopathy/dysplasia associated with woolly hair and epidermolytic palmoplantar keratoderma. The skin fragility-woolly hair syndrome (OMIM 607,655) represents another recessive disorder due to desmoplakin mutations characterized by palmoplantar keratoderma, woolly hair, variable alopecia, dystrophic nails, and excessive blistering. In these cases, as well as in the bovine case described here, there is no cardiomyopathy. Lethal acantholytic epidermolysis bullosa (OMIM 609,638), characterized by severe blistering of the epidermis and mucous membranes, is caused by homozygous deletions causing truncation of the DP tail [28]. Heterozygous carriers of any of these known recessive mutations displayed no phenotypic abnormalities in humans [29]. In general, epidermal fragility or excessive cornification is the first manifestation of these *DSP*-related human diseases, the hair changes (woolly hair or hypotrichosis) as well as the palmoplantar keratosis appear during childhood. The combination of these clinical features is an alarm sign for cardiomyopathy, which can appear at a young age in the form of severe arrhythmias, heart failure or spontaneous cardiac death [30]. In the *DSP*-associated case in cattle presented here, hyperkeratosis of the skin, changes in the hair coat as well as disorders of the mucous membranes are present. Therefore, at first glance, it resembles more the autosomal recessive forms of lethal acantholytic epidermolysis bullosa and/or skin fragility woolly hair syndrome than other *DSP*-related human diseases. Nevertheless, clear differences in manifestation are visible in detail, e.g. no evidence of blistering in the epidermis. Furthermore, there were no signs of heart problems, but a very young animal was examined and it cannot be ruled out that such changes would have occurred later in life. Also in human the classification of DSP-related phenotypes and the genotype-phenotype correlations are challenging, partly because different terms are used to designate disorders that comprise similar clinical features in the literature [306]. We therefore suspect that the specific *DSP* missense variant, which we identified here for the first time in a mammalian species, causes a highly probable unique syndromic disease phenotype. In most human cases, the precise consequences of the variants and the molecular pathology remain elusive due to the lack of expression and functional studies.

An exactly corresponding variant affecting human DSP residue 2288 was not yet described, and it not known as a variant in comprehensive databases such as gnomAD [31]. Interestingly, in a case report of a 14-year-old child with extensive mucocutaneous blisters caused by acantholysis of keratinocytes, epidermolytic palmoplantar keratoderma, nail dystrophy, enamel dysplasia, and sparse woolly hair, a very similar *DSP* missense variant was found [32]. This pathogenic variant results in a substitution of an aspartic acid for a conserved alanine residue at amino acid 2655 located in the C-terminal plakin domain of desmoplakin, affecting the third plakin-repeat subdomain of the human desmoplakin protein. The three tandem plakin repeat regions in the C-terminus of desmoplakin mediate binding to intermediate filaments. Its association with epidermal and simple keratins is dependent on the tertiary structure induced by heterodimerization of these intermediate filament proteins. Similar to the non-conservative amino acid replacement in the herein presented calf with acantholysis of the tongue epithelium, impaired desquamation of corneocytes in the skin (ichthyosis congenita), follicular / hair shaft dysplasia and corneal ulcers, such an exchange of a nonpolar, hydrophobic alanine with a charged acidic aspartate is most likely causative. In the previously cited human case description, it was reported that, although the variant does not significantly alter the three-dimensional structure of desmoplakin, structural modelling indicates changes in the electrostatic potential of the affected protein region [32]. The authors speculated that the change may seem subtle but the clinical phenotypes suggest that it alters intermediate filament binding functions that depend on electrostatically driven intermolecular interactions. Interestingly, immunofluorescence microscopy showed a reduction in the C-terminal domain of desmoplakin in the skin and oral mucosa of the child carrying this missense variant. Therefore, we conclude that the missense variant identified in the affected Scottish Highland calf represents a plausible candidate causative mutation for the lesions observed in this calf in the skin, the hair follicles, the tongue and the eyes. It is possible that the expression of the affected protein differs between the epithelium of the tongue and the epidermis, resulting in acantholysis in the tongue and impaired corneocyte desquamation in the skin. Impaired desmosome adhesion between the trichocytes of the hair shaft may also explain the alopecia and is supported by the histologically observed hair shaft dysplasia. Finally, the severe abnormalities of the cornea in the affected calf might be explained by the impaired function of desmoplakin as the corneal epithelium expresses a subgroup of keratins similar to those of epidermal epithelium [33]. Human genodermatoses, often have extracutaneous manifestations, and ocular manifestations in particular can have significant clinical implications, such as blindness [34]. The skin and eye malformations found in the affected calf resemble human keratitis-ichthyosis-deafness syndrome

(OMIM 148,210), a rare disorder caused by dominant acting variants in *GJB2* that encodes for connexin 26, a gap junction protein.

Conclusions

Rare disorders in livestock animals are traditionally poorly diagnosed. The report of this single case by a concerned breeder, followed by the diagnosis of ichthyosis congenita, follicular dysplasia, acantholysis of tongue epithelium, severe cornea defects, in combination with WGS has resulted in the identification of a most likely pathogenic variant in the *DSP* gene. This report should alert breeders of Scottish Highland cattle about the possible emergence of congenital ichthyosis in the future and will permit genetic testing to avoid the unintentional occurrence of further affected cattle. Screening the variant in the global Scottish Highland cattle population will enable better assessment of the population allele frequency for this breed-specific deleterious variant. Future studies to assess the functionality of the DSP protein in the presence of the missense variant will be valuable for understanding the biological impact of the mutation. To the best of our knowledge, this study represents the first report of a *DSP*-related syndromic form of congenital ichthyosis in domestic animals. The observed acantholysis of the tongue and corneal defects add *DSP* to the list of candidate genes for similar congenital phenotypes in humans.

Abbreviations

DSP: Desmoplakin; OMIA: Online Mendelian Inheritance in Animals; OMIM: Online Mendelian Inheritance in Man; WGS: Whole-genome sequencing.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12917-021-03113-3.

Additional file 1: Figure S1. Macroscopic pictures of the tongue of the affected calf after necropsy.

Additional file 2: Table S1. Private homozygous coding variants in the sequenced case. List of the remaining variants after the comparison to the control cohort of 705 genomes of other breeds, revealing 5 protein-changing variants only present in the genome of the ichthyosis-affected calf. These 5 variants with a moderate or high predicted impact on the encoded protein were located within 5 different genes or loci. Note that the predicted pathogenic variant NM_001192368.2: c.6893 C>A is the only one located in a functional candidate gene

Acknowledgements

We thank Nathalie Besuchet-Schmutz for expert technical assistance and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure.

Authors' contributions

AM, SR and MM conducted the clinical examinations. MMW performed the pathological examinations. CTK conducted pedigree analysis. IMH performed the bioinformatics of the whole-genome sequencing data and identified the most likely causative sequence variant. CD planned and supervised the genetic analysis. CTK performed sample collection of parents and controls. CTK and CD drafted the manuscript. All authors reviewed and edited during writing the manuscript. All authors have read and agreed to the final version of the manuscript.

Funding

This study was partially funded by the Swiss National Science Foundation, grant number 172911.

Availability of data and materials

The WGS data of the case can be found in the European Nucleotide Archive under the sample accession no. SAMEA5714970.

Declarations

Ethics approval and consent to participate

This study did not require official or institutional ethical approval as it was not experimental but part of clinical and pathological veterinary diagnostics. The study was carried out in compliance with the ARRIVE guidelines. The animal was handled according to good ethical standards and Swiss legislation (Animal Welfare regulation: Tierschutzverordnung from 23th April 2008, last amended on 4th September 2018). A formal ethics approval for the euthanasia of the affected calf was not needed as this complies with §179 of the Swiss Animal Welfare regulation. The affected calf was euthanized by intravenous injection of an overdose of sodium pentobarbital (300 mg/mL), performed by AM. The aim was to identify the cause of the congenital disorder. Written consent was obtained from the cattle owner.

Consent for publication

Not applicable.

Competing interests

The authors declare to have no financial or non-financial competing interests.

Author details

[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland. [2]Clinic for Ruminants, Vetsuisse Faculty, University of Bern, Bern, Switzerland. [3]Division of Clinical Dermatology, Vetsuisse Faculty, University of Bern, Bern, Switzerland. [4]Dermavet, Tierklinik Aarau-West, Oberentfelden, Switzerland. [5]Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland.

Received: 29 July 2021 Accepted: 13 December 2021
Published online: 07 January 2022

References

1. Marukian NV, Choate KA. Recent advances in understanding ichthyosis pathogenesis. F1000Res. 2016;5:F1000 Faculty Rev-1497.
2. Oji V, Tadini G, Akiyama M, Blanchet Bardon C, Bodemer C, Bourrat E, et al. Revised nomenclature and classification of inherited ichthyoses: results of the First Ichthyosis Consensus Conference in Sorèze 2009. J Am Acad Dermatol. 2010;63:607-41.
3. Vahlquist A, Fischer J, Törmä H. Inherited Nonsyndromic Ichthyoses: An Update on Pathophysiology, Diagnosis and Treatment. Am J Clin Dermatol. 2018;19:51–66.
4. Fischer J, Bourrat E. Genetics of Inherited Ichthyoses and Related Diseases. Acta Derm Venereol. 2020;100:adv00096.
5. Baker JR, Ward WR. Ichthyosis in domestic animals: a review of the literature and a case report. Br Vet J. 1985;141:1–8.
6. Julian RJ: Ichthyosis congenita in cattle. Vet Med. 1960;55:35–41.
7. Lüps P. Die Fischschuppenkrankheit (Ichthyosis universalis congenita), eine in Bayern beobachtete Erbkrankheit des Rindes. Berl Münch Tierärztl Wschr. 1963;76:204–6.
8. Raoofi A, Mardjanmehr SH, Nekoei S: Ichthyosis congenita in a calf in Iran. Vet Rec 2001;149:563.
9. Testoni S, Zappulli V, Gentile A. Ichthyosis in two Chianina calves. Dtsch Tierärztl Wochenschr. 2006;113:351–4.

10. Molteni L, Dardano S, Parma P, Polli M, De Giovanni AM, Sironi G, Longeri M. Ichthyosis in Chianina cattle. Vet Rec. 2006;158:412–4.
11. Cornillie P, Cools S, Vandaele L, de Kruif A, Simoens P. Ichthyosis in the Belgian white and red cattle breed. Vlaams Diergeneeskd Tijdschr 2007;76:345–51.
12. Cho JK, Son JM, Lee DS, Park SJ, Cho SW, Shin ST. Harlequin ichthyosis in a HanWoo calf. J Vet Med Sci. 2007;69:553–5.
13. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, et al: Highly effective SNP-based association mapping and management of recessive defects in livestock. Nat Genet. 2008;40:449–54.
14. Woolley SA, Eager KLM, Häfliger IM, Bauer A, Drögemüller C, Leeb T, O'Rourke BA, Tammen I. An ABCA12 missense variant in a Shorthorn calf with ichthyosis fetalis. Anim Genet. 2019;50:749–52.
15. Eager KLM, Conyers LE, Woolley SA, Tammen I, O'Rourke BA. A novel ABCA12 frameshift mutation segregates with ichthyosis fetalis in a Polled Hereford calf. Anim Genet. 2020;51:837–8.
16. Jacinto JGP, Häfliger IM, Veiga IMB, Letko A, Gentile A, Drögemüller C. A frameshift insertion in FA2H causes a recessively inherited form of ichthyosis congenita in Chianina cattle. Mol Genet Genomics. 2021;296:1313–22.
17. Bourneuf E, Otz P, Pausch H, Jagannathan V, Michot P, Grohs C, et al. Rapid Discovery of De Novo Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species. Sci Rep. 2017;7:11466.
18. Peters M, Reber I, Jagannathan V, Raddatz B, Wohlsein P, Drögemüller C. DNA-based diagnosis of rare diseases in veterinary medicine: a 4.4 kb deletion of ITGB4 is associated with epidermolysis bullosa in Charolais cattle. BMC Vet Res. 2015;11:48.
19. Jacinto JGP, Häfliger IM, Veiga IMB, Letko A, Benazzi C, Bolcato M, Drögemüller C. A Heterozygous Missense Variant in the COL5A2 in Holstein Cattle Resembling the Classical Ehlers–Danlos Syndrome. Animals. 2020;10:2002.
20. Jacinto JGP, Häfliger IM, Veiga IMB, Drögemüller C, Agerholm JS. A de novo mutation in KRT5 in a crossbred calf with epidermolysis bullosa simplex. J Vet Intern Med. 2020;34:2800–7.
21. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9:giaa021.
22. Häfliger IM, Sickinger M, Holsteg M, Raeder LM, Henrich M, Marquardt S, Drögemüller C, Lühken G. An IL17RA frameshift variant in a Holstein cattle family with psoriasis-like skin alterations and immunodeficiency. BMC Genet. 2020;21:55.
23. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. Annu Rev Anim Biosci. 2019;7:89–102.
24. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6:80–92.
25. Swiss Comparative Bovine Resequencing Project. https://www.ebi.ac.uk/ena/browser/view/PRJEB18113. (2021) Accessed 15 Jun 2021.
26. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31:2745–7.
27. Pigors M, Schwieger-Briel A, Cosgarea R, Diaconeasa A, Bruckner-Tuderman L, Fleck T, Has C. Desmoplakin mutations with palmoplantar keratoderma, woolly hair and cardiomyopathy. Acta Derm Venereol. 2015;95:337–40
28. Jonkman MF, Pasmooij AM, Pasmans SG, van den Berg MP, Ter Horst HJ, Timmer A, Pas HH. Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. Am J Hum Genet. 2005t;77:653–60.
29. Whittock NV, Wan H, Morley SM, Garzon MC, Kristal L, Hyde P, et al. Compound heterozygosity for non-sense and mis-sense mutations in desmoplakin underlies skin fragility/woolly hair syndrome. J Invest Dermatol. 2002;118:232–8.
30. Polivka L, Bodemer C, Hadj-Rabia S. Combination of palmoplantar keratoderma and hair shaft anomalies, the warning signal of severe arrhythmogenic cardiomyopathy: a systematic review on genetic desmosomal diseases. J Med Genet. 2016;53:289–95.
31. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME; Genome Aggregation Database Consortium, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581:434-43.
32. Mahoney MG, Sadowski S, Brennan D, Pikander P, Saukko P, Wahl J, et al. Compound heterozygous desmoplakin mutations result in a phenotype with a combination of myocardial, skin, hair, and enamel abnormalities. J Invest Dermatol. 2010;130:968–78
33. Kao WW. Keratin expression by corneal and limbal stem cells during development. Exp Eye Res. 2020;200:108206.
34. Jen M, Nallasamy S. Ocular manifestations of genetic skin disorders. Clin Dermatol. 2016;34:242–75.

Publisher's Note

A frameshift insertion in *FA2H* causes a recessively inherited form of ichthyosis congenita in Chianina cattle

Journal:	Molecular Genetics and Genomics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.1007/s00438-021-01824-8

Molecular Genetics and Genomics
https://doi.org/10.1007/s00438-021-01824-8

ORIGINAL ARTICLE

A frameshift insertion in *FA2H* causes a recessively inherited form of ichthyosis congenita in Chianina cattle

Joana G. P. Jacinto[1,2] · Irene M. Häfliger[2] · Inês M. B. Veiga[3] · Anna Letko[2] · Arcangelo Gentile[1] · Cord Drögemüller[2]

Received: 14 July 2021 / Accepted: 23 September 2021

Abstract

The aim of this study was to characterize the phenotype and to identify the genetic etiology of a syndromic form of ichthyosis congenita (IC) observed in Italian Chianina cattle and to estimate the prevalence of the deleterious allele in the population. Sporadic occurrence of different forms of ichthyosis including IC have been previously reported in cattle. However, so far, no causative genetic variant has been found for bovine IC. Nine affected cattle presenting congenital xerosis, hyperkeratosis and scaling of the skin as well as urolithiasis and cystitis associated with retarded growth were examined. Skin histopathology revealed a severe, diffuse orthokeratotic hyperkeratosis with mild to moderate epidermal hyperplasia. The pedigree records indicated a monogenic recessive trait. Homozygosity mapping and whole-genome sequencing allowed the identification of a homozygous frameshift 1 bp insertion in the *FA2H* gene (c.9dupC; p.Ala4ArgfsTer142) located in a 1.92 Mb shared identical-by-descent region on chromosome 18 present in all cases, while the parents were heterozygous as expected for obligate carriers. These findings enable the selection against this sub-lethal allele showing an estimated frequency of ~7.5% in Chianina top sires. A sporadic incidence of mild clinical signs in the skin of heterozygous carriers was observed. So far, pathogenic variants affecting the encoded fatty acid 2-hydroxylase catalyzing the synthesis of 2-hydroxysphingolipids have been associated with myelin disorders. In conclusion, this study represents the first report of an *FA2H*-related autosomal recessive inherited skin disorder in a mammalian species and adds *FA2H* to the list of candidate genes for ichthyosis in humans and animals. Furthermore, this study provides a DNA-based diagnostic test that enables selection against the identified pathogenic variant in the Chianina cattle population. However, functional studies are needed to better understand the expression of FA2H in IC-affected Chianina cattle.

Keywords Bovine · Genodermatoses · Fatty acid 2-hydroxylase · Precision medicine · Skin · Urolithiasis

Introduction

The aim of this study was to report a series of nine cases of IC in Chianina cattle, to characterize the clinicopathological phenotype and finally to present the results of the genetic

Communicated by Joan Cerdá.

✉ Cord Drögemüller
cord.droegemueller@vetsuisse.unibe.ch

Joana G. P. Jacinto
joana.goncalves2@studio.unibo.it

Irene M. Häfliger
irene.haefliger@vetsuisse.unibe.ch

Inês M. B. Veiga
ines.veiga@vetsuisse.unibe.ch

Anna Letko
anna.letko@vetsuisse.unibe.ch

Arcangelo Gentile
arcangelo.gentile@unibo.it

[1] Department of Veterinary Medical Sciences, University of Bologna, Ozzano Emilia, 40064 Bologna, Italy

[2] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland

[3] Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland

Published online: 02 October 2021

analysis that evidenced a homozygous frameshift variant in bovine *FA2H* gene. Moreover, the prevalence of the deleterious allele in a selected population of Chianina sires is also estimated.

Genodermatoses are sporadic inherited disorders of the skin that both in humans and in livestock animals mostly follow a monogenic mode of inheritance (Leeb et al. 2017; Jacinto et al. 2020; Pope 2020).

In human medicine, the concept of 'genodermatosis with skin fragility' was recently introduced by adding other genetic disorders with skin fragility such erosive or hyperkeratotic disorders (Pope 2020). Inherited ichthyosis characterized by an abnormal terminal keratinocyte differentiation belongs to this group of skin fragility disorders encompassing a clinically, pathologically and heritably heterogeneous presentation with a thickened stratum corneum resulting in localized or generalized scaling (Marukian and Choate 2016).

In human medicine, the classification of the different forms of ichthyosis is based on clinicopathological manifestations and mode of inheritance, being divided in two main types: non-syndromic forms when clinical findings are limited to the skin, and syndromic forms in case additional organs are involved (Oji et al. 2010). In this respect, ichthyosis has been associated to pathogenic variants in more than 30 genes that are involved in several cellular functions, such as DNA repair, lipid biosynthesis, adhesion and desquamation (Oji et al. 2010). In particular, recent advances have reinforced the causative role of mutations in genes encoding proteins essential to the formation of the hydrophobic barrier (Marukian and Choate 2016).

In domestic animals, ichthyosis has been described in dogs (Credille et al. 2009; Grall et al. 2012; Metzger et al. 2015; Bauer et al. 2017; Casal et al. 2017), pigs (Wang et al. 2019), sheep (Câmara et al. 2017) and cattle (Charlier et al. 2008; Woolley et al. 2019; Eager et al. 2020). Furthermore, it has also been reported in greater kudu calves (Chittick et al. 2002). While in sheep, pigs and greater kudu the underlying genetic cause of this condition has not been determined, in dogs pathogenic variants have been identified in five different candidate genes associated with the phenotype, four breed specific recessive inherited forms (*TGM1*, *SLC27A4*, *PNPLA1*, *NIPAL4*) (OMIA 000546-9615; OMIA 001973-9615; OMIA 001588-9615; OMIA 001980-9615) as well as a single dominant inherited form in a single affected dog (*ASPRV1*) (OMIA 002099-9615).

In Chianina, Shorthorn and Polled Hereford cattle, a form of ichthyosis named ichthyosis fetalis, which resembles the Harlequin-type ichthyosis described in human medicine, has been associated with recessively inherited mutations in *ABCA12* (OMIA 002238-9913) (Charlier et al. 2008; Woolley et al. 2019; Eager et al. 2020). Affected calves are stillborn or die within the first days after birth and the skin is diffusely covered with large horny plates separated by deep fissures and resembling a 'leather cuirass'. Furthermore, eversion at mucocutaneous junctions provokes eclabium and ectropion (Chittick et al. 2002; Molteni et al. 2006).

In Chianina cattle, a second less severe form of ichthyosis, named ichthyosis congenita (IC), has also been described (Testoni et al. 2006) and in subsequent time repeatedly presented to the authors. Animals with IC show milder but comparable lesions to those of ichthyosis fetalis. It is clinically characterized by a more or less extended scale-like hyperkeratosis and multifocal alopecic areas, and histopathologically by a diffuse lamellar orthokeratotic hyperkeratosis. The underlying genetic cause of this form of syndromic form of ichthyosis associated with retarded growth is unknown.

Methods

Animals

This study did not require official or institutional ethical approval as it was not experimental, but rather part of clinical and pathological veterinary diagnostics. All animals in this study were examined with the consent of their owners and handled according to good ethical standards. It deals with a total of 129 Chianina cattle, including 9 IC-affected animals, 4 dams, 6 sires and 110 artificial insemination (AI) top sires. The tenth affected animal included in the study (case 10) was the one previously reported by Testoni et al. (2006), whose blood had at that time been frozen and therefore had remained available for genetic studies.

Clinical and pathological investigations

Eight calves (cases 1–8) and one heifer (case 9) presenting cutaneous hyperkeratosis and retarded growth were recorded by the teaching hospital of the Department of Veterinary Medical Sciences, University of Bologna between 2005 and 2020 (Online Resource 1). The mean age of record of the calves was 2.6 months (minimum–maximum: 2 days–7 months), whereas the heifer was 18 months. The mean age at death was 11.2 months (natural death, euthanasia or slaughtering). All affected animals and one dam (case 8's dam) were thoroughly clinically examined. Information related to the skin condition of the other dams as well as of the sires were obtained by interviewing the owners or the breeders' association, respectively.

A parasitological test for detection of ectoparasites and fungi infection was performed on case 8's dam.

Skin biopsies using an 8 mm punch were obtained from seven affected animals (cases 1–3 and cases 6, 8, 9) and from case 8's dam. The collected samples were fixed in 10% neutral buffered formalin, trimmed, processed, embedded in

paraffin wax, sectioned at 4 µm, and stained with hematoxylin and eosin (H&E) for further histological evaluation. Two affected animals (cases 2 and 8) were submitted to necropsy.

Pedigree design

Pedigree analysis was performed using Pedigraph version 2.4 software (Department of Animal Science, University of Minnesota, USA).

DNA extractions

Genomic DNA was extracted from the IC-affected animals (EDTA blood samples), four related dams (EDTA blood samples) and six related sires (semen) using Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland). Furthermore, genomic DNA was also obtained from semen of 110 Chianina AI top sires with the same methodology.

SNP array genotyping and homozygosity mapping

High-density SNP genotyping was carried out for seven cases (cases 1–7) and eight obligate carriers (three dams and five sires) (Online Resource 1) on the Illumina BovineHD BeadChip array including 777.962 SNPs. All given SNP positions correspond to the bovine ARS-UCD1.2 genome assembly. The PLINK v1.9 software (Chang et al. 2015) was used to perform basic quality filtering of the dataset. Even though no sample was excluded, a total of 146.440 variants were removed owing to minor allele thresholds. The total genotyping rate was approximately 0.98. With a total of 631.522 remaining markers, homozygosity mapping was performed for the 7 IC-affected animals using the software PLINK v1.9 (Purcell et al. 2007) with the commands *--homozyg-kb* 100 (considering homozygous segments of at least 100 kb), *--homozyg-match* 0.95 (for allelic matching between both cases) and *--homozyg-group* (for generating an overlap-file), resulting in shared runs of homozygosity (ROH) indicating chromosomal region of identity-by-descent (IBD).

Whole-genome sequencing and variant calling

WGS using the Illumina NovaSeq6000 (Illumina Inc., San Diego, CA, USA) was performed on the genomic DNA of two affected calves (cases 1 and 6). The sequenced reads were mapped to the ARS-UCD1.2 reference genome, resulting in an average read depth of approximately 18.2× in case 1 and 17.9× in case 6, and single-nucleotide variants (SNVs) and small indel variants were called (Rosen et al. 2020). The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format files were in accordance with the 1000 Bull Genomes Project processing guidelines of run 7 (Hayes and Daetwyler 2019), except for the trimming, which was performed using fastp (Chen et al. 2018). Further preparation of the genomic data was done according to Häfliger et al. 2020 (Häfliger et al. 2020). To find private variants, we compared the genotypes of the two calves with 597 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that are publicly available (Online Resource 2) in the European Nucleotide Archive (SAMEA7690197 and SAMEA7690198 are the samples accession number of case 1 and case 6, respectively; http://www.ebi.ac.uk/en). The filtered list of remaining variants were further checked for their occurrence in a global control cohort of 4110 genomes of a variety of breeds (Hayes and Daetwyler 2019). Integrative Genomics Viewer (IGV) (Robinson et al. 2017) software was used for visual inspection of genome regions containing possible candidate genes.

Variant validation and genotyping via Sanger sequencing

PCR and Sanger sequencing were used to confirm the WGS results and to perform targeted genotyping for the identified *FA2H* frameshift insertion variant (18:2205625C>CG). All IC-affected animals, four available dams and six sires, as well as 113 AI top sires that included three fathers of the studied cases, were genotyped for the identified variant. Also the case reported by Testoni in 2006 (case 10) (Testoni et al. 2006) was genotyped. Primers were designed using the Primer-BLAST tool (Ye et al. 2012). After amplification with AmpliTaqGold360Mastermix (Thermo Fisher Scientific) the purified PCR products were directly sequenced on an ABI3730 capillary sequencer (Thermo Fisher Scientific). The primer sequences used were the following: 5′-AAATTCCTGGTT-GGGGAGCC-3′ (forward primer) and 5′-CTCGACAACGAGACGCACC-3′ (reverse primer). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes).

Results

Clinical phenotype

All patients (case 1–9) showed a more or less extended skin xerosis, hyperkeratosis and scaling besides a retarded growth. In the affected area the skin was dry and greyish with scale-like hyper-keratosis, and the most severe lesions were present at the level of trunk and neck (Fig. 1a). The coat was dull and bristly. Moreover, multifocal alopecic lesions were noticed, mostly affecting the muzzle, eyelids, ears, and inner region of limbs. The youngest calves displayed multiple wrinkles and folds (≤ 1 month of age)

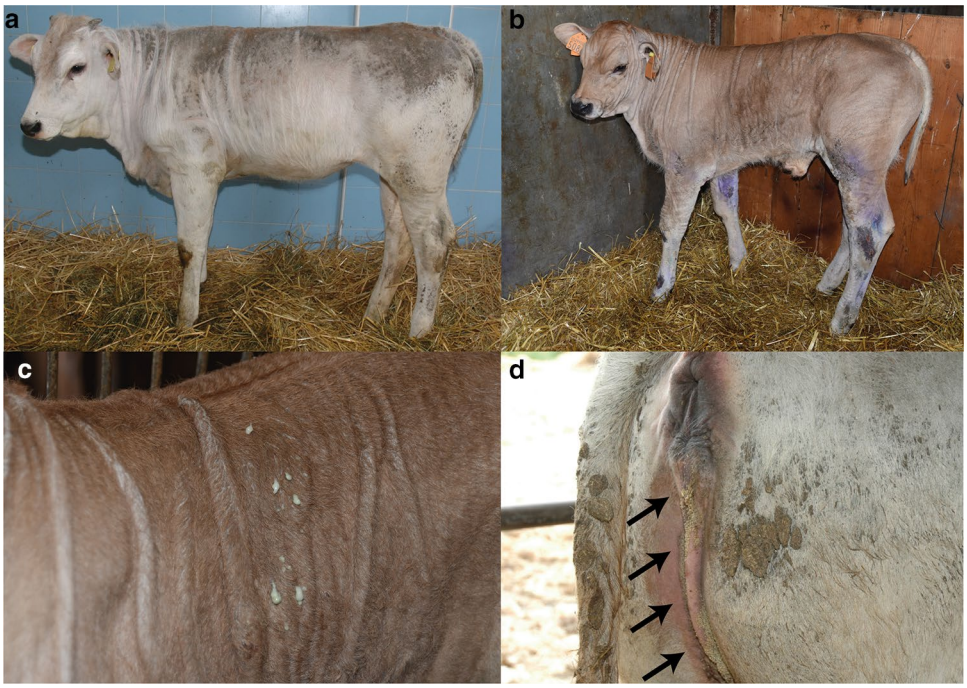


Fig. 1 Clinical characterization of Chianina cattle affected by ichthyosis congenita. **a** Note the dry, greyish skin with scale-like hyperkeratosis over most of the body surface (case 1). **b** Note the multiple wrinkles, folds and wounds secondary to the hyperkeratosis (case 8). **c** Higher magnification of **b** from the skin of the thoracic region. Note the pyoderma. **d** Note the urolithiasis characterized by the presence of small stones and crystals (arrows) on the perigenital region (case 7)

(Fig. 1b). No abnormalities were observed at the level of the mucocutaneous junctions. One of the animals (case 8) also showed secondary wounds and pyodermitis (Fig. 1c). Urolithiasis evidenced by the presence of small stones and crystals in the perigenital region (Fig. 1d) accompanied the cutaneous disease in cases 1, 7, 8 and 10. A hypoglycemic and hypothermic crisis that provoked the death of case 8 during the winter season was interpreted as a secondary phenomenon of imbalanced thermoregulation capacity. No abnormalities were registered at the level of the cardiovascular, respiratory, musculoskeletal, and nervous systems in any animals.

Interestingly, case 8's dam showed mild localized xerosis, hyperkeratosis and scaling in the region of the rump. Unfortunately, since in most cases the parents of the affected animals had already been slaughtered, we could not evaluate the phenotype in more of these obligate carriers. However, we did see a total of three other confirmed *FA2H* heterozygous Chianina cattle and they were clinically normal.

Based on the clinical observations, the affected animals were consequently suspected to suffer from IC as described in this breed in 2006 (Testoni et al. 2006). A similar diagnosis was advanced also for the dam of case 8, although in a very mild form. For this animal the differential diagnosis of ectoparasitosis and fungi infection were excluded on the base of a parasitological test.

Pathological phenotype

Histological analysis of the biopsies from the cutaneous lesions revealed a severe, diffuse orthokeratotic hyperkeratosis with mild to moderate epidermal hyperplasia (Fig. 2a, b). Serocellular crusts, serum lakes, and plant material were occasionally present among the abundant keratin scales. The superficial dermis displayed multifocal, moderate eosinophilic infiltrates, as well as a mostly perivascular, moderate infiltration with plasma cells and lymphocytes (Fig. 2a). Also, it was possible to observe the presence of intracytoplasmic, spindle-shaped, optically empty clefts within the sebocytes (Fig. 2c) in several of the affected animals, while the remaining adnexal structures were unremarkable. Similar findings were observed histologically in the punch biopsies taken from the dam of case 8 (Fig. 2d). These findings were consistent with the clinical diagnosis of IC.

Moreover, post-mortem examination of three cases (cases 1, 8 and 10) revealed inflammation of the urinary bladder (cystitis).

Genetic analysis

Pedigree records allowed the identification of a common ancestor as all IC-affected Chianina cattle were inbred from a sire born in 1976 (Online Resource 3). Pedigree analysis was consistent with monogenic autosomal recessive inheritance, and therefore carried out homozygosity mapping as all cases would likely be homozygous for a common chromosome segment flanking the causal mutation. This revealed a total of two identical-by-descent (IBD) segments shared by all seven cases with available SNP data (case 1–7): one 548 kb-sized region on

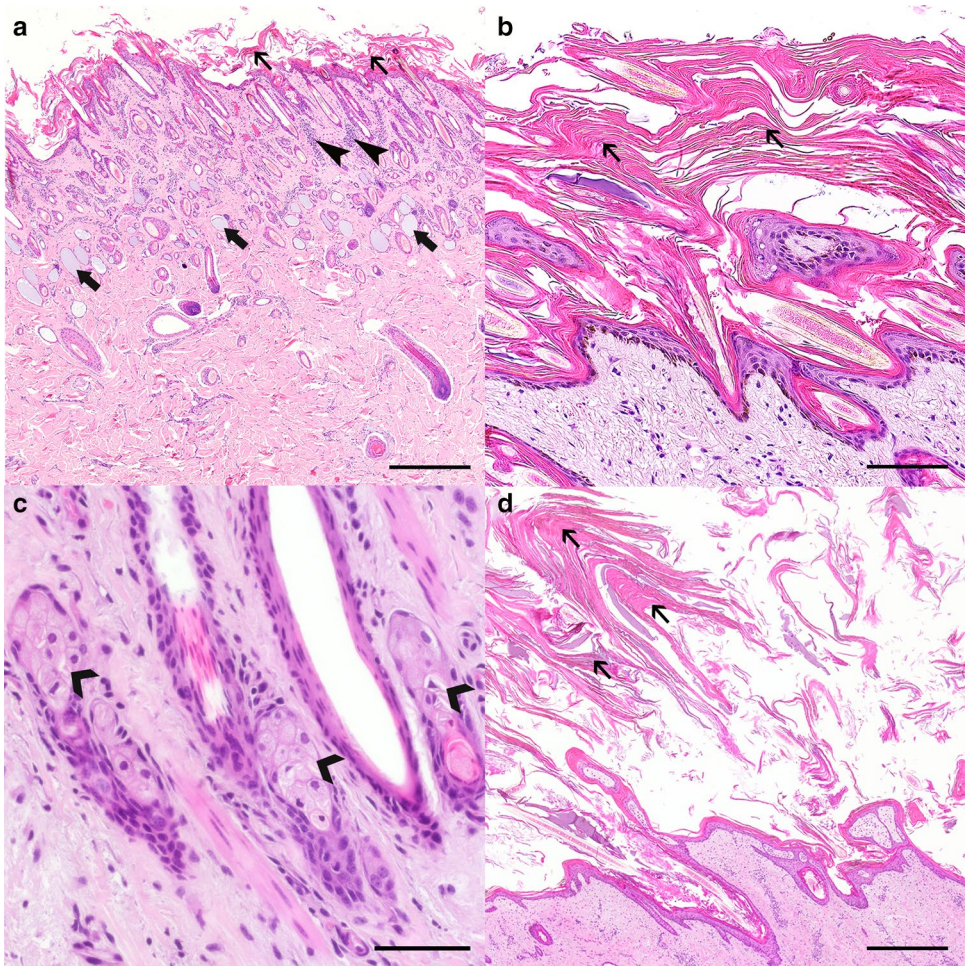


Fig. 2 Histology of the skin lesions displayed by a IC-affected Chianina calf (**a**–**c**) and its dam (**d**). **a** The epidermis of the calf (case 8) is irregular and mildly hyperplastic, with a thick overlying stratum corneum composed of abundant orthokeratotic, lamellar keratin scales (thin arrows). The sebaceous glands are not noticeable at this magnification, and the sweat glands are often dilated and filled with basophilic, homogeneous material (large arrows). Occasional interstitial inflammatory infiltrates can be observed in the superficial dermis (thin arrowheads). H&E staining, 500 µm. **b** Detail of the abundant orthokeratotic, lamellar keratin scales (thin arrows) (case 8). H&E staining, 100 µm. **c** Detail of the sebaceous glands. Some sebocytes display intracytoplasmic, spindle-shaped, optically empty clefts (large arrowheads) (case 8). H&E staining, 50 µm. **d** A severe orthokeratotic hyperkeratosis (thin arrows) could also be observed in the dam from case 8. H&E staining, 500 µm

chromosome 5 from 21.75 to 22.298 Mb and a second 1.92 Mb-sized region on chromosome 18 from 1.37 to 3.29 Mb (Fig. 3a).

Filtering of WGS for private shared homozygous variants present in sequenced genomes of cases 1 and 6 and absent in 597 available control genomes identified ten private protein-changing variants with a predicted moderate or high impact. Analyzing the occurrence of these variants in the global control cohort of 4110 genomes of a variety of breeds (Hayes and Daetwyler 2019), a single frameshift variant in *FA2H* with a predicted high impact on the encoded protein exclusively present in the genome of the case 1 and 6 remained. The homozygous variant in *FA2H* exon 1 on chromosome 18 (chr18:g.2205625C>CG; c.9dupC) was confirmed using IGV software (Fig. 3b, d). The deleterious *FA2H* variant (NM_001192455.1: c.9dupC) was predicted to result in a frameshift in the beginning of the protein after alanine 4 with a stop codon after aspartate 142 (NP_001179384.1: p.Ala4ArgfsTer142) resulting in a completely different amino acid sequence, if expressed, when compared with the wild type protein (Fig. 3e, f).

FA2H genotyping

To confirm and evaluate the presence of the *FA2H* variant, the affected genomic region was amplified by PCR and Sanger sequenced (Fig. 3c) in a total of ten cases, and presumable dams and sires when available. Analyzing the sequencing data, we observed that all cases were homozygous for the variants, whereas the available parents were heterozygous as expected for obligate carriers (Table 1). Furthermore, genotyping of 113 Chianina sires representing the active breeding population revealed a carrier ratio of 15% whereas the variant was absent in a global cohort of more than 4700 cattle of various breeds (Table 1).

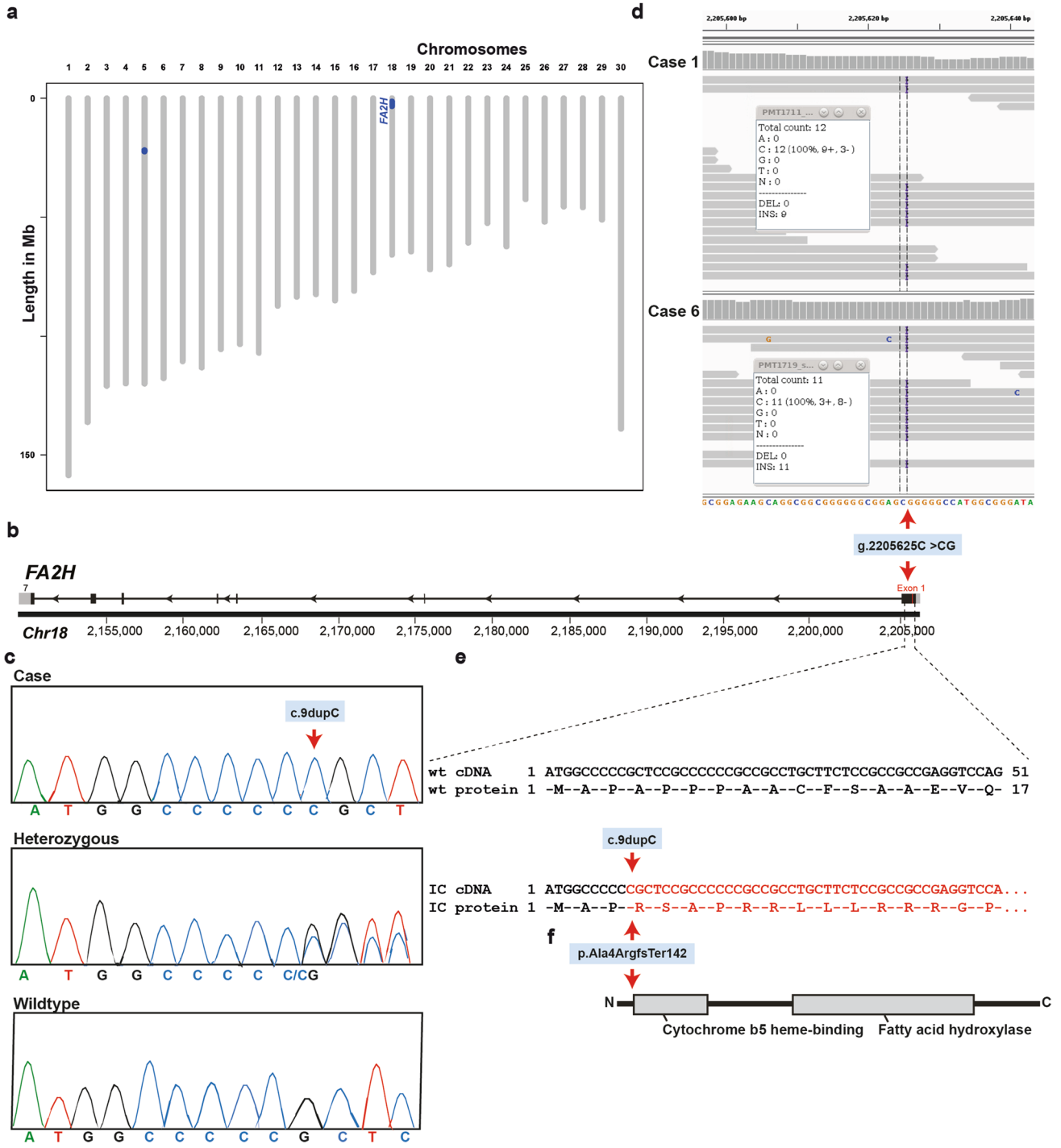


Fig. 3 Ichthyosis congenita (IC) *FA2H* frameshift variant in Chianina cattle. **a** Genetic mapping of the IC locus in the cattle genome. The two regions of shared homozygosity of seven cases are displayed in blue. Note that the largest segment of 1.92 Mb on chromosome 18 containing the, *FA2H* gene. **b** *FA2H* gene structure showing the variant located in exon 1. **c** Electropherograms of a case, heterozygous and wild-type genotypes. **d** Integrative Genomics Viewer (IGV) screenshot presenting the g.2205625C>CG variant in the two whole-genome sequenced cases. **e** Predicted wild-type (wt) and IC cDNA e protein. **f** Schematic representation of the bovine FA2H protein and its two domains and the identified pathogenic frameshift variant (p.Ala4fsTer142; red arrow)

Table 1 Association of the 1 bp duplication (c.9dupC) variant in *FA2H* with ichthyosis congenita (IC) in Chianina cattle

	Genotype		
	wt/wt	wt/dup	dup/dup
IC-affected cattle	0	0	10
Obligate carriers[a]	0	10	0
Chianina top sires	96	17	0
Normal control cattle from various breeds	4707	0	0

[a]Parents of the affected animal

Discussion

Here we describe the clinicopathological phenotype displayed by nine Italian Chianina cattle with IC and present the results of the genetic analysis that identified a recessively inherited frameshift mutation in *FA2H*, providing a novel candidate gene for skin disorders in both humans and animals. Furthermore, we provide a DNA-based diagnostic test that enables the selection against this sub-lethal allele that show an estimated frequency of ~7.5% in Chianina top sires.

Clinicopathological resemblances between familial forms of ichthyosis in humans and animals and the keratinization defect observed histologically in IC-affected Italian Chianina cattle led to the hypothesis that genetic variants in candidate genes for ichthyosis could be responsible for this disease in cattle. However, protein-changing variants within these more than 30 known candidate genes (Oji et al. 2010) were not found within the two mapped IBD regions, thereby excluding these as likely candidates. We then performed whole-genome sequencing of two IC-affected cattle that led to the identification of a frameshift mutation in *FA2H* (c.9dupC; p.Ala4ArgfsTer142). The genetic association of this variant with the bovine familial IC phenotype was confirmed by the homozygous genotype in eight additionally affected cattle, including an older case presented in 2006, and by its absence in all other sequenced genomes. Furthermore, reported expression of *FA2H* transcripts in the urinary tract supports the associated urolithiasis and cystitis seen in some of the IC-affected cattle. Finally, the predicted consequence of the frameshift variant demonstrating a loss-of-function supports causality. This variant is therefore the first in any domestic animal species to be associated with IC, and *FA2H* should be considered an additional candidate gene for syndromic forms of ichthyosis in humans.

While genetic analysis strongly suggested the association of p.Ala4ArgfsTer142 allele with IC in affected Italian Chianina cattle, the frameshift variant lies very near the N-terminal end of the protein. Therefore, the impact of such a significant truncation on protein function is probably high, thus the variant represents a most likely pathogenic loss-of-function mutation. Within a representative cohort of the current Italian Chianina population, a moderate allele frequency and the absence of the homozygous genotype for the deleterious allele was noticed.

In humans, mutations in *FA2H* (OMIM611026) are associated with recessively inherited spastic paraplegia type 35 (Dick et al. 2010), leukodystrophy with spasticity and dystonia (Edvardson et al. 2008), and fatty acid hydroxylase-associated neurodegeneration, a rare subtype neurodegeneration with brain iron accumulation (Kruer et al. 2010). So far, more than 40 different mutations have been associated with these neurological phenotypes (Rattay et al. 2019; Kawaguchi et al. 2020). However, to the best of our knowledge, no pathogenic variant in the *FA2H* associated to a form of ichthyosis has been reported in both animal and human species. Therefore, our study in cattle provides the first large-animal model of an *FA2H*-related congenital skin disorder.

FA2H encodes the endoplasmic reticulum enzyme fatty acid 2-hydroxylase, which plays a major role in the de novo synthesis of sphingolipids containing 2-hydroxy fatty acids (Alderson et al. 2004, 2005; Maldonado et al. 2008). 2-Hydroxy sphingolipids are very plentiful in neural tissue since the major components of myelin are galactolipids (galactosylceramide and sulfatide) with 2-hydroxy fatty acids (Alderson et al. 2005; Maldonado et al. 2008). However, the function of 2-hydroxyl modification of sphingolipids is still poorly known, although several studies evidently demonstrated that these compounds (including ceramides) play important roles in signal transduction (Hannun and Obeid 2008). Particularly, a study demonstrated that absence of FA2H lead to the impairment of cAMP-dependent cell cycle exit of Schwannoma cells, suggesting that FA2H sphingolipids may negatively regulate the cell cycle (Alderson and Hama 2009). Moreover, *FA2H* is highly expressed in the epidermis (Uchida et al. 2007). Notably, mammalian skin contains reasonably large amounts of 2-hydroxylated sphingolipids, which are involved in cell–cell recognition, signal transduction, and intercellular adhesion (Hakomori 2002; Uchida et al. 2007). The sphingolipids' ceramide backbone also plays a role as an intracellular signal of cell arrest, cellular senescence, and apoptosis in several types of cell, including keratinocytes (Hannun and Luberto 2000). Besides these ubiquitous bioregulatory functions, ceramide are abundant components of the extracellular lamellar membranes in the outermost layers of the epidermis, such as the stratum corneum, where they play an important role in the epidermal permeability barrier function (Holleran et al. 2006). Also, a notable increase in ceramide is noticed during epidermal differentiation (Holleran et al. 2006). In humans, it is known that differentiation-dependent up-regulation of ceramide synthesis and fatty acid elongation is accompanied by up-regulation of FA2H. Furthermore, the 2-hydroxylation of fatty acid by FA2H occurs prior to generation of

ceramides/glucosylceramides, and 2-hydroxyceramides/2-hydroxyglucosylceramides are essential for epidermal lamellar membrane formation (Uchida et al. 2007). Such findings suggest that the expression of FA2H is essential for epidermal permeability barrier homeostasis and responsible for synthesis of 2-hydroxylated sphingolipids in keratinocytes of mammalian skin. Therefore, the severe orthokeratotic hyperkeratosis observed in the Chianina cattle with IC could be a consequence of the frameshift insertion in the *FA2H* gene. Notably, lack of FA2H in *Fa2h*$^{-/-}$ mice leads to hyperproliferation of sebocytes and enlarged sebaceous glands during hair follicle morphogenesis and anagen (active growth phase) in adult mice (Maier et al. 2011). Interestingly, the IC-affected animals included in this study often displayed intracytoplasmic, spindle-shaped, optically empty clefts within the sebocytes, although the sebaceous glands were similar in size to the ones observed in control animals. Sebaceous glands are holocrine glands that secrete a viscous, lipid-rich fluid rich in cholesterol and wax esters, triglycerides, squalene and cholesterol playing an important role in thermoregulation (Porter 2001). The rate of sebum is associated with the number and size of glands, and low production of sebum might lead to sebatosis or xerosis (Porter 2001; Shamloul and Khachemoune 2020). The major functions of sebum are to lubricate the skin and hair conferring impermeability to water, and in thermoregulation (Shamloul and Khachemoune 2020). Moreover, sebaceous glands play a role in immunity since sebum is thought to have antibacterial and antifungal properties (Strauss et al. 1983). Consequently, our findings may suggest that cattle affected by IC might be predisposed to develop skin secondary infections and present thermoregulation deficits due to this genetic defect. These situations were evidently suspected for two of our patients. Moreover, in *Fa2h*$^{-/-}$ mice, deficiency in *Fa2h* caused a delay in emergence of the fur during morphogenesis and depilation-induced anagen and a cyclic alopecia (Maier et al. 2011). Herein, the cases revealed a localized alopecia, but the hair follicles present in the biopsies taken from the IC cutaneous lesions were unremarkable.

In mice, depletion of FA2H decreases the protein levels of GLUT4 leading to reduced glucose uptake and lipogenesis under basal and insulin-stimulated conditions (Guo et al. 2010). GLUT4 deficiency in mice (Slc2a4^{tm1Mch}/Slc2a4^{tm1Mch}) leads to retarded growth, decreased expected longevity and abnormal cellular glucose and fat metabolism (Katz et al. 1995). Intriguingly, all Chianina cattle homozygous for the *FA2H* mutation showed retarded growth and decreased expected longevity. Unfortunately, metabolic analysis to access the glucose metabolism was not performed.

Beside the skin lesions that were displayed in all the cases, four out of ten cases showed urolithiasis and three out of ten revealed cystitis. It is worth to highlight that in the six animals where these findings were not recorded: one was euthanized 2 days after birth, and consequently, the absence of these findings might be explain by the young age of the calf; the remaining five were clinically examined only at the farms and, therefore, it was not possible to have a follow-up of the clinical status. On the contrary, the cases where we observed urolithiasis and cystitis were recovered at the clinic allowing the performance of a more detailed examination. Urolithiasis is a multifactorial disease resulting from complex interactions between environmental and genetic factors. Interestingly, *FA2H* in humans is also expressed in the urinary bladder and kidney (Fagerberg et al. 2014). However, physiological functions of FA2H in these organs are largely unknown.

Conclusions

Rare disorders like IC in livestock are usually not reported or are mis-diagnosed. Based on the known function of *FA2H* and its role in the skin, the predicted impact of the identified variant and its perfect co-segregation with the disease phenotype in the studied pedigree, we conclude that inherited IC in Chianina cattle is caused by a homozygous loss-of-function variant in *FA2H*. Thereby, this study represents an outstanding animal model for the understanding of similar conditions in different species and adds *FA2H* to the list of candidate genes for ichthyosis in humans. This example highlights the utility of precision diagnostics including genomics, for understanding rare disorders and the value of surveillance of cattle breeding populations for harmful genetic disorders.

Moreover, this study provides a DNA-based diagnostic test that allows selection against the identified pathogenic variant in the Chianina cattle population. Due to the high economic value of many Chianina cattle, including their skin for leather production, genetic testing should be pursued to prevent breeding of carriers that produce affected calves.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00438-021-01824-8.

Acknowledgements The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the WGS experiments and providing high-performance computing infrastructure. The technical assistance of Nathalie Besuchet Schmutz, Manuela Bozzo, Erika Bürgi, and Francesca Caravello is acknowledged.

Author contributions JGPJ performed all clinical and genetic analyses and drafted the manuscript. IMH carried out bioinformatics. AL assisted in the genetic analyses. AG collected samples and drafted parts of the manuscript. IMBV performed histopathology and drafted parts of the manuscript. AG and CD designed the study, supervised the project and finalized the manuscript. All authors participated in writing the manuscript and have read and approved the final version.

Funding Open Access funding provided by Universität Bern. This study was partially funded by the Swiss National Science Foundation.

Data availability The whole-genome data of our group have been made freely available under study accession number PRJEB28191in the European Nucleotide Archive (http://www.ebi.ac.uk/ena). All accession numbers of the WGS are available in the Online Resource 2. SAMEA7690197 and SAMEA7690198 are the samples accession number of case 1 and case 6, respectively. All references to the bovine FA2H gene correspond to the NCBI accessions NC_037345.1 (chromosome 18, ARS-UCD1.2), NM_001192455.1 (*FA2H* gene), and NP_001179384.1 (FA2H protein). For the protein structure of FA2H, the UniProt database accession number E1BGC2 was used.

Code availability Not applicable.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethics approval This study did not require official or institutional ethical approval as 'non-experimental clinical veterinary practices' are specifically excluded from being considered regulated procedures under The Animals (Scientific Procedures) Act (ASPA), 1986, Section 2(8) (https://www.rcvs.org.uk/setting-standards/advice-and-guidance/code-of-professional-conduct-for-veterinary-surgeons/supporting-guidance/recognised-veterinary-practice/). The cattle were handled according to good ethical standards and all live animals were blood sampled by a veterinary for diagnostic purposes on the farm to determine the cause of the disease. All other sampling was carried out postmortem on affected animals after euthanasia on humane grounds. The aim was to identify the cause of the congenital disorder and thereby improve the animal welfare situation on the Chianina cattle population by identifying the underlying genetic cause and preventing breeding of further cases.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Alderson NL, Hama H (2009) Fatty acid 2-hydroxylase regulates cAMP-induced cell cycle exit in D6P2T Schwannoma cells. J Lipid Res 50:1203–1208. https://doi.org/10.1194/jlr.M800666-JLR200

Alderson NL, Rembiesa BM, Walla MD et al (2004) The human FA2H gene encodes a fatty acid 2-hydroxylase. J Biol Chem 279:48562–48568. https://doi.org/10.1074/jbc.M406649200

Alderson NL, Walla MD, Hama H (2005) A novel method for the measurement of in vitro fatty acid 2-hydroxylase activity by gas chromatography–mass spectrometry. J Lipid Res 46:1569–1575. https://doi.org/10.1194/jlr.D500013-JLR200

Bauer A, Waluk DP, Galichet A et al (2017) A de novo variant in the ASPRV1 gene in a dog with ichthyosis. PLoS Genet 13:1–13. https://doi.org/10.1371/journal.pgen.1006651

Câmara ACL, Borges PAC, Paiva SA et al (2017) Ichthyosis fetalis in a cross-bred lamb. Vet Dermatol 28:125–516. https://doi.org/10.1111/vde.12459

Casal ML, Wang P, Mauldin EA et al (2017) A Defect in NIPAL4 is associated with autosomal recessive congenital ichthyosis in American bulldogs. PLoS ONE 12:1–9. https://doi.org/10.1371/journal.pone.0170708

Chang CC, Chow CC, Tellier LCAM et al (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4:1–16. https://doi.org/10.1186/s13742-015-0047-8

Charlier C, Coppieters W, Rollin F et al (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. Nat Genet 40:449–454. https://doi.org/10.1038/ng.96

Chen S, Zhou Y, Chen Y, Gu J (2018) Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chittick EJ, Olivry T, Dalldorf F et al (2002) Harlequin ichthyosis in two greater Kudu (*Tragelaphus strepsiceros*). Vet Pathol 39:751–756. https://doi.org/10.1354/vp.39-6-751

Credille KM, Minor JS, Barnhart KF et al (2009) Transglutaminase 1-deficient recessive lamellar ichthyosis associated with a LINE-1 insertion in Jack Russell terrier dogs. Br J Dermatol 161:265–272. https://doi.org/10.1111/j.1365-2133.2009.09161.x

Dick KJ, Eckhardt M, Paisán-Ruiz C et al (2010) Mutation of FA2H underlies a complicated form of hereditary spastic paraplegia (SPG35). Hum Mutat 31:1251–1260. https://doi.org/10.1002/humu.21205

Eager KLM, Conyers LE, Woolley SA et al (2020) A novel ABCA12 frameshift mutation segregates with ichthyosis fetalis in a Polled Hereford calf. Anim Genet 51:837–838. https://doi.org/10.1111/age.12973

Edvardson S, Hama H, Shaag A et al (2008) Mutations in the fatty acid 2-hydroxylase gene are associated with leukodystrophy with spastic paraparesis and dystonia. Am J Hum Genet 83:643–648. https://doi.org/10.1016/j.ajhg.2008.10.010

Fagerberg L, Hallström BM, Oksvold P et al (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteom 13:397–406

Grall A, Guaguère E, Planchais S et al (2012) PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nat Genet 44:140–147. https://doi.org/10.1038/ng.1056

Guo L, Zhou D, Pryse KM et al (2010) Fatty acid 2-hydroxylase mediates diffusional mobility of raft-associated lipids, GLUT4 level, and lipogenesis in 3T3-L1 adipocytes. J Biol Chem 285:25438–25447. https://doi.org/10.1074/jbc.M110.119933

Häfliger IM, Wiedemar N, Švara T et al (2020) Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. Anim Genet 51:382–390. https://doi.org/10.1111/age.12923

Hakomori SI (2002) The glycosynapse. Proc Natl Acad Sci USA 99:225–232. https://doi.org/10.1073/pnas.012540899

Hannun YA, Luberto C (2000) Ceramide in the eukaryotic stress response. Trends Cell Biol 10:73–80. https://doi.org/10.1080/00397911.2014.916302

Hannun YA, Obeid LM (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. Nat Rev Mol Cell Biol 9:139–150. https://doi.org/10.1038/nrm2329

Hayes BJ, Daetwyler HD (2019) 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. Annu Rev Anim Biosci 7:89–102. https://doi.org/10.1146/annurev-animal-020518-115024

Holleran WM, Takagi Y, Uchida Y (2006) Epidermal sphingolipids: metabolism, function, and roles in skin disorders. FEBS Lett 580:5456–5466. https://doi.org/10.1016/j.febslet.2006.08.039

Jacinto JGP, Häfliger IM, Veiga IMB et al (2020) A heterozygous missense variant in the COL5A2 in Holstein cattle resembling the classical Ehlers–Danlos syndrome. Animals 10:1–14. https://doi.org/10.3390/ani10112002

Katz EB, Stenbit AE, Hatton K, DePinho R, Charron MJ (1995) Cardiac and adipose tissue abnormalities but not diabetes in mice deficient in GLUT4. Nature 377(6545):151–155

Kawaguchi M, Sassa T, Kidokoro H et al (2020) Novel biallelic FA2H mutations in a Japanese boy with fatty acid hydroxylase-associated neurodegeneration. Brain Dev 42:217–221. https://doi.org/10.1016/j.braindev.2019.11.006

Kruer MC, Paisán-Ruiz C, Boddaert N et al (2010) Defective FA2H leads to a novel form of neurodegeneration with brain iron accumulation (NBIA). Ann Neurol 68:611–618. https://doi.org/10.1002/ana.22122

Leeb T, Müller EJ, Roosje P, Welle M (2017) Genetic testing in veterinary dermatology. Vet Dermatol 28:4-e1. https://doi.org/10.1111/vde.12309

Maier H, Meixner M, Hartmann D et al (2011) Normal fur development and sebum production depends on fatty acid 2-hydroxylase expression in sebaceous glands. J Biol Chem 286:25922–25934. https://doi.org/10.1074/jbc.M111.231977

Maldonado EN, Alderson NL, Monje PV et al (2008) FA2H is responsible for the formation of 2-hydroxy galactolipidsin peripheral nervous system myelin. J Lipid Res 49:153–161. https://doi.org/10.1194/jlr.M700400-JLR200

Marukian NV, Choate KA (2016) Recent advances in understanding ichthyosis pathogenesis [version 1; referees: 2 approved]. F1000Research 5:1–9

Metzger J, Wöhlke A, Mischke R et al (2015) A novel SLC27A4 splice acceptor site mutation in Great Danes with ichthyosis. PLoS ONE 10:1–13. https://doi.org/10.1371/journal.pone.0141514

Molteni L, Dardano S, Parma P et al (2006) Ichthyosis in Chianina cattle. Vet Rec 158:412–414. https://doi.org/10.1136/vr.158.12.412

Oji V, Tadini G, Akiyama M et al (2010) Revised nomenclature and classification of inherited ichthyoses: results of the First Ichthyosis Consensus Conference in Sorze 2009. J Am Acad Dermatol 63:607–641. https://doi.org/10.1016/j.jaad.2009.11.020

Pope E (2020) Epidermolysis bullosa: a 2020 perspective. Br J Dermatol 183:603. https://doi.org/10.1111/bjd.19125

Porter AMW (2001) Why do we have apocrine and sebaceous glands? J R Soc Med 94:236–237. https://doi.org/10.1177/014107680109400509

Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. https://doi.org/10.1086/519795

Rattay TW, Lindig T, Baets J et al (2019) FAHN/SPG35: a narrow phenotypic spectrum across disease classifications. Brain 142:1561–1572. https://doi.org/10.1093/brain/awz102

Robinson JT, Thorvaldsdóttir H, Wenger AM et al (2017) Variant review with the integrative genomics viewer. Cancer Res 77:e31–e34. https://doi.org/10.1158/0008-5472.CAN-17-0337

Rosen BD, Bickhart DM, Schnabel RD et al (2020) De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 9:1–9. https://doi.org/10.1093/gigascience/giaa021

Shamloul G, Khachemoune A (2020) An updated review of the sebaceous gland and its role in health and diseases Part 1: embryology, evolution, structure, and function of sebaceous glands. Dermatol Ther. https://doi.org/10.1111/dth.14695

Strauss J, Downing D, Ebling F (1983) Sebaceous glands. In: Goldsmith L (ed) Biochemistry and physiology of the skin. Oxford University Press, New York, pp 569–595

Testoni S, Zappulli V, Gentile A (2006) Ichthyosis in two Chianina calves. Dtsch Tierarztl Wochenschr 113:351–354

Uchida Y, Hama H, Alderson NL et al (2007) Fatty acid 2-hydroxylase, encoded by FA2H, accounts for differentiation-associated increase in 2-OH ceramides during keratinocyte differentiation. J Biol Chem 282:13211–13219. https://doi.org/10.1074/jbc.M611562200

Wang X, Cao C, Li Y et al (2019) A harlequin ichthyosis pig model with a novel ABCA12 mutation can be rescued by acitretin treatment. J Mol Cell Biol 11:1029–1041. https://doi.org/10.1093/jmcb/mjz021

Woolley SA, Eager KLM, Häfliger IM et al (2019) An ABCA12 missense variant in a Shorthorn calf with ichthyosis fetalis. Anim Genet 50:749–752. https://doi.org/10.1111/age.12856

Ye J, Coulouris G, Zaretskaya I et al (2012). BMC Bioinform. https://doi.org/10.1186/1471-2105-13-134

X-linked hypohidrotic ectodermal dysplasia in crossbred beef cattle due to a large deletion in *EDA*

Journal:	Animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, performing genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.3390/ani11030657

animals

MDPI

Article

X-Linked Hypohidrotic Ectodermal Dysplasia in Crossbred Beef Cattle Due to a Large Deletion in *EDA*

Donal O'Toole [1,*], Irene M. Häfliger [2], Fabienne Leuthard [2,3], Brant Schumaker [1], Lynn Steadman [4], Brian Murphy [5], Cord Drögemüller [2,†] and Tosso Leeb [2,3,†]

1 Wyoming State Veterinary Laboratory, University of Wyoming, Laramie, WY 82070, USA; bschumak@uwyo.edu
2 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch (I.M.H.); fabileuthard@gmail.com (F.L.); cord.droegemueller@vetsuisse.unibe.ch (C.D.); tosso.leeb@vetsuisse.unibe.ch (T.L.)
3 Dermfocus, University of Bern, 3001 Bern, Switzerland
4 Chadron Veterinary Clinic, Chadron, NE 69337, USA; lynns@panhandle.net
5 Department of Pathology, Microbiology and Immunology, School of Veterinary Medicine, University of California, Davis, CA 95616, USA; bmurphy@ucdavis.edu
* Correspondence: DOT@uwyo.edu; Tel.: +1-307-766-9976
† These authors contributed equally to the work.

Simple Summary: Ectodermal dysplasias such as hypohidrotic ectodermal dysplasia (HED), are genetic conditions affecting the development and/or homeostasis of two or more ectodermal derivatives, including hair, teeth, nails, and eccrine glands. In particular, X-linked hypohidrotic ectodermal dysplasia-1 (ECTD1) in humans is characterized by a triad of signs comprising sparse hair, abnormal teeth, and anhidrosis or hypohidrosis. It has been reported in cattle, dogs, mice and rats. Until now, eight pathogenic variants in the bovine ectodysplasin A (*EDA*) gene causing ECTD1-like disorders have been found. Herein, five affected Red Angus-Simmental bull calves born over a 6-year period (2013–2019) in a single herd in the Western United States are reported showing an ECTD1-like syndrome. Calves were born with severe hypotrichosis and oligodontia. At age 1-week-old two calves died of severe pneumonia. Microscopic findings of the skin revealed small-caliber hair follicles with a mean density in flank skin slightly greater in affected animals than in control animals. Nasolabial, intranasal and tracheobronchial mucosal glands were absent, whereas olfactory glands were unaffected. Whole-genome sequencing (WGS) identified a 53 kb deletion of the X chromosome including parts of the *EDA* gene as well as the entire *AWAT2* gene. The partial deletion of the *EDA* gene that is known to be associated with forms of ECTD1 is the most likely cause for the reported genodermatosis. Similar rare disorders in livestock are often not diagnosed at the molecular level due to lack of resources, short lifespan of the animals, and concerns for the producers' reputation.

Abstract: X-linked hypohidrotic ectodermal dysplasia-1 (ECTD1) in people results in a spectrum of abnormalities, most importantly hypotrichosis, anodontia/oligodontia, and absent or defective ectodermally derived glands. Five Red Angus-Simmental calves born over a 6-year period demonstrated severe hypotrichosis and were diagnosed as affected with ECTD1-like syndrome. Two died of severe pneumonia within a week of birth. The skin of three affected calves revealed a predominance of histologically unremarkable small-caliber hair follicles. Larger follicles (>50 µm) containing medullated hairs (including guard and tactile hairs) were largely restricted to the muzzle, chin, tail, eyelids, tragus and distal portions of the limbs and tail. The mean histological density of hair follicles in flank skin of two affected calves was slightly greater than that in two unaffected calves. One affected calf was examined postmortem at 10 days of age to better characterize systemic lesions. Nasolabial, intranasal and tracheobronchial mucosal glands were absent, whereas olfactory glands were unaffected. Mandibular incisor teeth were absent. Premolar teeth were unerupted and widely spaced. Other than oligodontia, histological changes in teeth were modest, featuring multifocal disorganization of ameloblasts, new bone formation in dental alveoli, and small aggregates of osteodentin and cementum at the margins of the enamel organ. A 52,780 base pair deletion spanning six out of

Citation: O'Toole, D.; Häfliger, I.M.; Leuthard, F.; Schumaker, B.; Steadman, L.; Murphy, B.; Drögemüller, C.; Leeb, T. X-Linked Hypohidrotic Ectodermal Dysplasia in Crossbred Beef Cattle Due to a Large Deletion in *EDA*. *Animals* **2021**, *11*, 657. https://doi.org/10.3390/ani11030657

Academic Editors: Francesca Ciotola, Sara Albarella and Vincenzo Peretti

Received: 29 January 2021
Accepted: 25 February 2021
Published: 2 March 2021

eight coding exons of *EDA* and all of *AWAT2* was identified. Partial deletion of the *EDA* gene is the presumed basis for the reported X-chromosomal recessive inherited genodermatosis.

Keywords: *Bos taurus*; cattle; ectodysplasin; hypodontia; hypotrichosis; oligodontia; whole genome sequencing; development; dermatology; genodermatosis

1. Introduction

Ectodermal dysplasias are inherited diseases in which development and/or homeostasis of two or more ectodermal derivatives such as hair, teeth, sweat and other glands, and occasionally nails, are abnormal [1]. The most common ectodermal dysplasia in people is hypohidrotic ectodermal dysplasia (HED) [2]. Initially classified by clinical presentation and mode of inheritance, members of the HED group are categorized by genotype and molecular pathway, in addition to phenotype [2]. X-linked hypohidrotic ectodermal dysplasia-1 (ECTD1) (OMIM 305100) is caused by variants in *EDA*, which encodes the transmembrane protein ectodysplasin A (EDA), a member of the TNFα-related signaling pathway [3]. Charles Darwin was among the first to describe what later proved to be ECTD1 in a kindred of toothless men from Sind, noting clinically relevant features such as oligodontia, hairlessness and heat intolerance [4]. More than 200 pathogenic variants in human *EDA* are recognized. Affected individuals suffer from heat intolerance due to a limited ability to sweat. Respiratory infections are common in some human forms of ECTD1 and are attributed to absent and/or reduced numbers of glands in the respiratory tract [5]. Additional clinical problems are dental disease, eczema and xerostomia. Similar clinical forms of HED also occur. These are due to pathogenec variants in the *EDAR* gene encoding the EDA receptor, variants in the *EDARADD* gene encoding the EDAR-associated death domain, and. variants in the *WNT10A* gene encoding the secreted signaling molecule Wnt family member 10A. As a result, there are different autosomal recessive and dominant forms of HED [1]. Various forms of HED can be clinically and pathologically indistinguishable because one signal transduction pathway is involved [3].

HED has been documented in mice, dogs, cattle and rats [6–12]. Our understanding of EDA-related anhidrotic ectodermal dysplasia is based in large part on the *Tabby* (Eda^{Ta}) mouse [11,13–15]. The ability to ameliorate or abolish congenital defects of HED using short-term treatment with recombinant ectodysplasin was first demonstrated in *Tabby* mice and refined in a canine model, leading to successful treatment of human fetuses [16–19]. To date, there are eight reported pathogenic variants of the bovine *EDA* gene that, result in an ECTD1-like disease (OMIA 000543-9913) [8–10,20–27]. They are potentially lethal since affected cattle are ill-thrifty due to oligodontia and the heightened risk for pneumonia and hypothermia [20,21,27]. Affected cattle can be kept alive when fed chopped feed and in circumstances where there is limited exposure to extreme weather conditions [22,24,27]. Additionally, an HED syndrome affecting Charolais cattle due to a genetic variant in the *EDAR* gene was reported. Similar to its human counterpart, the bovine *EDAR*-related condition is transmitted as an autosomal recessive trait (OMIA 002128-9913) [28]. There are accounts in the peer-reviewed literature of other HED-like syndromes which are based on clinical signs and lesions alone [29–32].

Published reports of HED in domestic animals focus on the causative genetic variant so that information about lesions other than hypotrichosis and oligodontia is limited. Studies of affected people, mice and dogs document a pattern of lesions in addition to those in skin and teeth, particularly in glands of the oropharynx, ear and orbit, which contribute to clinical and subclinical disease [13,33–35]. Information about histological changes in teeth and respiratory tract of ECTD1-affected cattle is limited [24]. Gross changes in the teeth of affected people and dogs include characteristic conical crowns, hypodontia and/or oligodontia, malocclusion, delayed eruption, persisting deciduous

teeth, missing permanent teeth, decreased number of cusps, and decreased or diminutive tooth roots [36–39]. Comparable dental anomalies affect *Tabby* mice [14]

The aim of this study was to characterize the major systemic lesions in a small group of affected Red Angus-Simmental cattle born over a 6-year period in a single herd, and to identify the causative genetic variant.

2. Materials and Methods

2.1. Herd history and Clinical Investigation

Five crossbred bull calves were born with hypotrichosis and oligodontia over a 6-year period (2013–2019) in a commercial Red Angus-Simmental herd in western Nebraska, USA. Both natural breeding and artificial insemination were used. Each calf was born to a unique dam and was full-term based on the owner's assessment of body size and birth weight. The owner changed ear tags once he decided to retain animals in the herd. No key was kept for the two sets of identification numbers so it was not possible to generate a pedigree of the dams. Each of the five dams was clinically normal, according to the owner, although no examination of the oral cavity was performed to check for abnormal dentition in four of the five cattle. Skin biopsies were collected from three affected calves (cases 1 and 2 in 2013; case 3 in 2019; cases 4 and 5 were assessed by clinical examination alone) by a veterinary practitioner (L.S.) in order to establish a diagnosis for alopecia. After case 3 was born in 2019, the owner examined the incisors of this fifth dam, a first-calf heifer, and found no abnormalities such as missing or malformed teeth. The dam's cheek teeth (premolars and molars) were not examined. Case 3 was kept indoors from birth in a heated barn, and wrapped in an insulated coat to prevent pneumonia. Once it was established from a skin biopsy that the calf was free of bovine viral diarrhea virus (a recognized cause of congenital alopecia due to in utero infection) and that histological lesions in skin were indistinguishable to those in cases 1 and 2, the calf was humanely euthanized for diagnostic purposes at 10 days after birth. No affected calves were born subsequently into the herd in 2020. As of this writing, a single affected male calf was born into the herd in Feb 2021.

2.2. Pathological Investigation

Skin samples (5 × 5 cm) obtained by biopsy from the three affected calves were fixed in 10% buffered formalin. Skin was sampled from one pinna and the dorsal aspect of the neck (case 1), pinna alone (case 2), and pinna, muzzle, flank, upper and lower eyelids, lateral aspect of thigh, distal aspect of forelimbs and hindlimbs, tailhead, and distal aspect of the tail (case 3). Skin samples from calf 3 were selected to include lightly haired areas (e.g., flank) and better haired areas (e.g., distal aspect of limbs and tail; eyelids). A comprehensive set of tissues was collected into 10% formalin from case 3 at necropsy. These included tissues from oral and nasal cavities (tongue; cheek with oral mucosa; hard palate; nasal and ethmoidal conchae), palatine tonsils, larynx, trachea, bronchi, and lung including bronchioles. No samples were collected to evaluate lacrimal, ceruminous or major salivary glands. Maxillae and mandibles of case 3 were radiographed after death, fixed for 7 days in 10% buffered formalin, cut with a band saw along the long axis of teeth, and examined grossly. Teeth, mandibles and maxillae were decalcified in a formic acid-EDTA solution (CalFor*TM, Cancer Diagnostics Inc., Durham, NC, USA). After fixation, tissues were routinely processed through graded alcohols before embedding in paraffin. Sections were cut 5 µm thick and stained with hematoxylin and eosin (HE). Sections of skin from calves 1 and 3 were cut either perpendicular or parallel to the epidermal surface. The diameter of hair follicles was determined by stereology using flank skin from two affected (cases 1 and 3) and two age- and herd-matched Red Angus-Simmental control calves. The cross-sectional area of follicles areas was determined using a 4-point nucleator tool and converted to diameters (Stereo Investigator Software, MBF Bioscience, Williston, VT, USA). The density of hair follicles was measured using commercial image-capturing software (Olympus cellSens, Tokyo, Japan). Sections of aural skin from cases 1–3 were stained immunohistochemically for bovine viral diarrhea virus (BVDV) antigen using

appropriate positive and negative controls performed as described by Cornish et al. [40]. For tissues controls, samples of flank skin were collected postmortem from two female Red Angus-Simmental 5- or 6-day-old calves from the same ranch where affected calves were born. Death was due to neonatal disease, based on field necropsies performed by one author (L.S.). One calf had acute pneumonia and the other necrotizing hemorrhagic enteritis. The skull of a phenotypically normal 8-day-old Galloway bull calf that died of myocardial necrosis and interstitial pneumonia served as a control for the histology of normal bovine teeth and mucosa of nasal and ethmoidal conchae. All three control calves had normal pelage and dentition.

2.3. Genetic Investigation

Genomic DNA was isolated from the spleen of case 3 using a commercial kit (MagMax; Applied Biosystems/Thermo Fisher, Waltham, MA, USA). An Illumina Truseq PCR-free fragment library of the DNA from case 3 was prepared, generating ~176 million 2 × 150 bp read-pairs corresponding to 17.6× coverage on a NovaSeq 6000 instrument (illumina, San Diego, CA, USA). The reads were mapped to the ARS-UCD1.2 cattle reference genome (GCF_002263795.1) and subsequently single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format files were in accordance with the 1000 Bull Genomes Project processing guidelines (www.1000bullgenomes.com (accessed on 11 February 2021)). Further preparation of the genomic data had been done according to Häfliger et al., 2020 [41]. The Integrative Genomics Viewer (IGV) software was used for visual inspection and identification of structural variants [42]. The whole-genome sequencing (WGS) data are available at the European Nucleotide Archive under study accession PRJEB18113 and sample accession SAMEA5714973 (https://www.ebi.ac.uk/ena/data/view/ERS3518602 (accessed on 11 February 2021)). In order to find private variants, we compared the genotypes of the affected calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies at the Institute of Genetics at the Vetsuisse Faculty, University of Bern, and that are publicly available in the European Nucleotide Archive under study accession PRJEB18113 (http://www.ebi.ac.uk/en (accessed on 11 February 2021)).

3. Results

3.1. Clinicopathological Phenotype

The five affected calves appeared almost completely hairless at birth, with the exception of eyelashes, muzzle, chin, tragus, and distal parts of the limbs and tail (Figure 1A). Vibrissae were present on the muzzle and chin. Small fine hairs suggestive of undercoat (pili lanei) were evident over the trunk, neck, head and upper two thirds of the limbs (Figure 1B). Digits were unremarkable, including periople and horn.

The histological appearance of hair follicles in the affected calves were indistinguishable from the smallest follicles in the control calves. The majority of hair follicles (>90%) contained lightly pigmented hairs and had a normal relationship with unremarkable arrector pili muscles and sebaceous glands (Figure 1C). Apocrine glands were sac-like with light blue contents in HE-stained sections and were lined by low cuboidal epithelium. They lay deep to hair follicles and were closely associated with the dermal papillae. Depending on the area of skin sampled from calf 3, small hair follicles were either the sole or predominant follicle type present (Figure 1D). This animal, at 10 days the oldest of the three affected calves whose skin was examined histologically, was the only one to have mild diffuse perivascular lymphocytic dermatitis.

The median diameter of hair follicles in flank skin was 36.4 and 38.4 µm in affected calves 1 and 3, compared to 51.1 and 51.3 µm in two control calves (Figure 2). The density of hair follicles in flank skin of the same two affected calves was slightly greater than that of two calves with normal hair coats (Table 1).

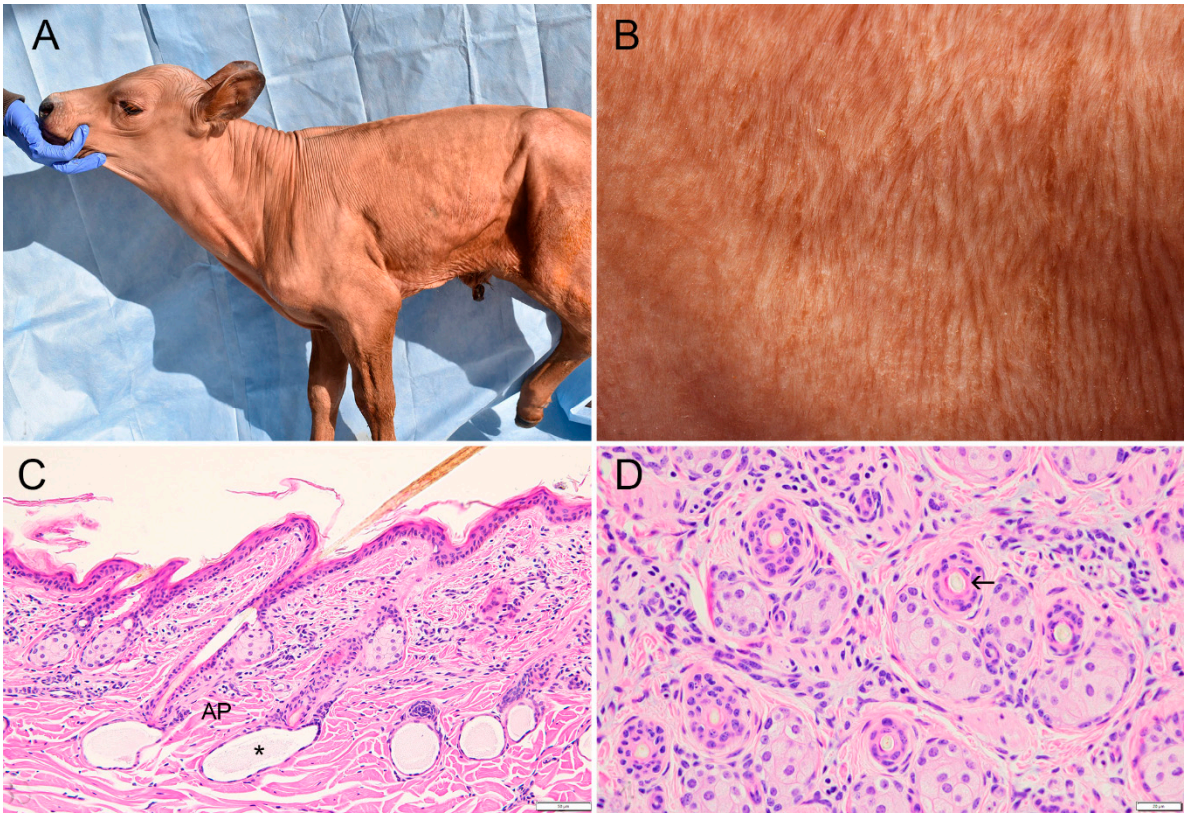


Figure 1. Cutaneous clinicopathological findings in a Red Angus-Simmental calf with hypohidrotic ectodermal dysplasia. (**A**) Case 3 at 10 days of age. Absence of visible hairs on neck, trunk, upper limbs or most of the face, with exception of eyelashes and pinna. (**B**) Case 3 at 10 days of age. Higher magnification of skin from flank showing small fine hairs grouped in linear dorso-ventral patterns. (**C**) Flank skin; case 3. Uniformly small hair follicles with associated sebaceous glands, pigmented hair shafts, and arrector pili muscles (AP). Sweat glands are closely associated with dermal papillae and contain lightly stained secretion (asterisk). Note mild lymphocytic dermatitis in upper to mid-dermis. Bar: 50 µm. (hematoxylin and eosin stain (H&E)). (**D**) Flank skin; case 1. Follicles contain hair shafts (arrow) and are associated with unremarkable sebaceous glands. No orphan sebaceous glands are present. Bar: 25 µm. (H&E).

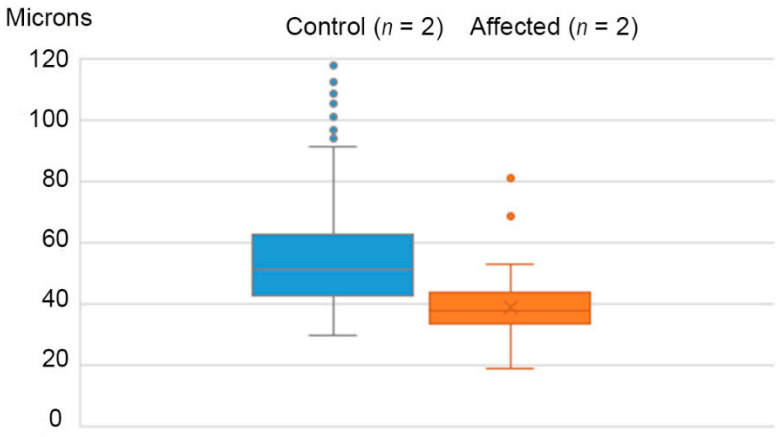


Figure 2. Comparison of median diameter of hair follicles in two control calves (blue) and two affected calves (orange; cases 1 and 3). In the control calves, 14 follicles were >120 µm diameter and not illustrated; they are included in the median. Median follicular diameters are 49.9 and 38.1 µm in unaffected and affected animals, respectively.

Table 1. Density of hair follicles in flank skin of affected and unaffected calves.

Calves		Number of Hair Follicles/Area (153,137 μm^2) [a]			Mean Follicle Density/Area (153,137 μm^2)	Mean Density Follicles/mm^2
		Area 1	Area 2	Area 3		
Affected	No. 1	30	31	30	30.3	164.50
Affected [b]	No. 3	22	18	23	21.0	
Unaffected [c]	No. 1	20	23	22	21.7	133.9
Unaffected [d]	No. 2	18	21	19	19.3	

[a]: Measured in skin from flank in all 4 calves. [b]: Mild diffuse lymphocytic dermatitis in this calf may have slightly reduced follicle density. [c]: Died at 5 days of age of pneumonia; normal hair coat. [d]: Died at 6 days of age of necrotizing enteritis; normal hair coat.

Normal Meibomian glands were present in the eyelids; no effort was made to quantify them. Nasolabial glands were absent in the planum of calf 3. The skin of cases 1–3 was negative for BVDV by immunohistochemistry.

Case 3 had 12 deciduous teeth, all of which were unerupted or barely erupted premolars (oligodontia). Normal neonatal calves have 20 deciduous teeth (incisors, canines and premolars) [43]. The dental formula in case 3 was deciduous incisors 0/0, deciduous canines 0/0, deciduous premolars 3/3, and permanent molars 0/0. In normal healthy calves, the crown of the permanent first mandibular molars is generally grossly evident at birth [43]. By contrast, in case 3 by 10 days of age no permanent teeth were evident radiographically, grossly or histologically. Unerupted or barely erupted premolars were slender, slightly curved, and more widely spaced than the closely packed premolars of an 8-day old calf used as a control (Figure 3A). Microscopically, premolar teeth demonstrated all components of normally developing bovine teeth. No dental matrices (dentin/enamel) of incisor, canine or molar teeth were detected histologically. In spite of gross malformation in the shapes of the premolar teeth, histological changes were modest. They comprised focal disorganization of ameloblast palisades, deposition of enamel-like aggregates in stratum intermedium (Figure 3B), disorderly spurs of dentin or osteodentin with cementum (Figure 3C), and limited new bone formation in dental alveoli (Figure 3D). Normal minor salivary glands were present in mucosa of the cheek and palate. Olfactory glands were present in mucosa of the ethmoidal conchae (Figure 3E). By contrast, no tracheal or bronchial-bronchiolar glands were present in the respiratory tract (Figure 3F,G). There was no evidence of inflammation in the upper or lower respiratory tract of case 3.

3.2. Genetic Analysis

Filtering of the obtained variant catalogue for private variants exclusively present in the genome of the affected calf and absent in 496 available control genomes identified any private protein-changing variants. As our automated variant calling was limited to single nucleotide and small indel variants, we additionally performed a visual analysis for large structural variants in four known HED candidate genes (*EDA, EDAR, EDARADD, WNT10A*). This analysis revealed a large ~53 kb deletion on the X chromosome in the affected calf 3. It can be formally designated as ChrX:80,382,423_80,435,202del52,780 (ARS-UCD1.2 assembly). The deletion was flanked by highly similar sequences that showed 95% sequence identity over 259 bp suggesting that the deletion may have arisen during an unequal crossing over event. The deleted region encompassed the entire *AWAT2* gene and the last six exons of the *EDA* gene (Figure 4).

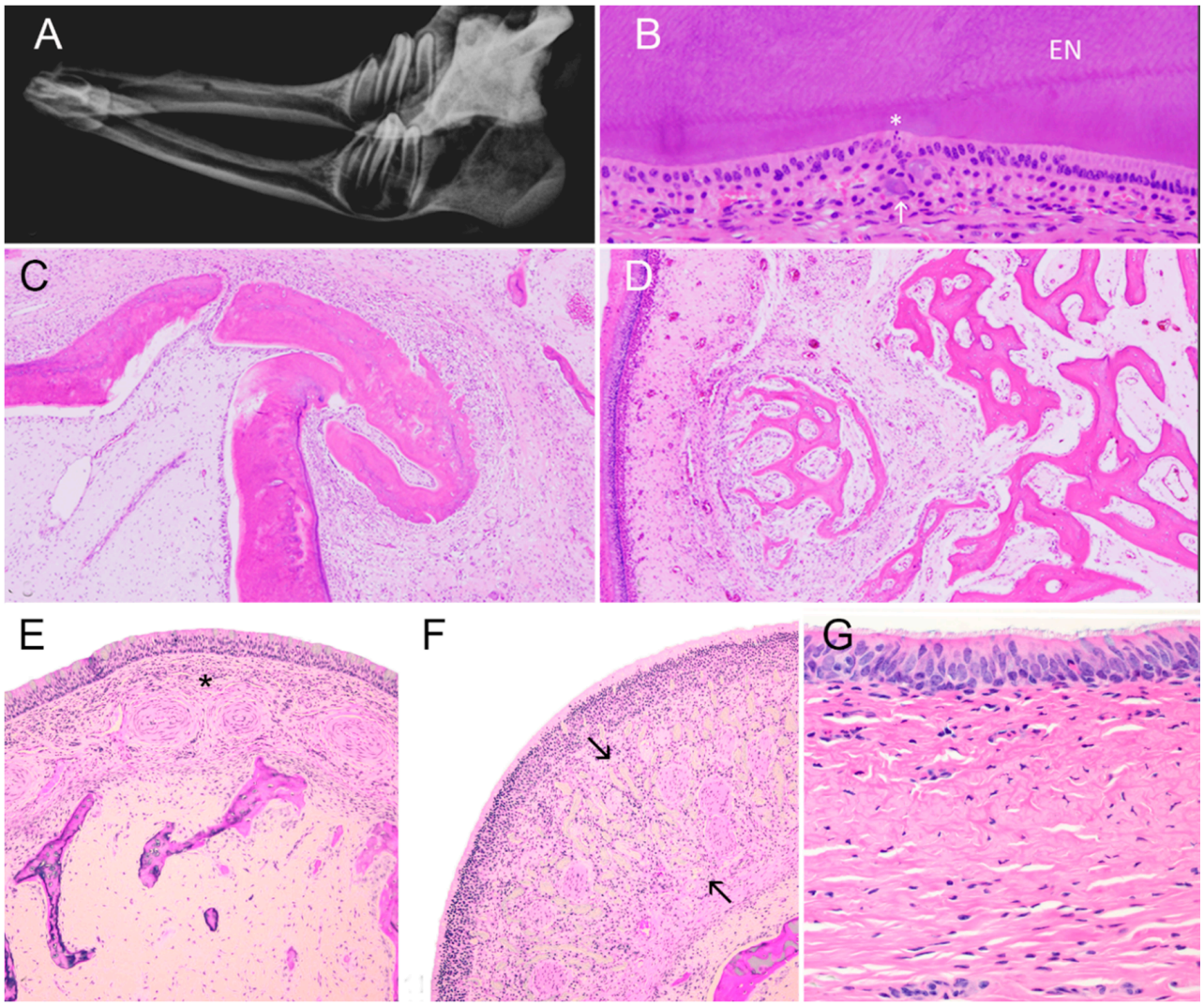


Figure 3. Dental anomalies in a hypohidrotic ectodermal dysplasia-affected calf. (**A**) Mandible (radiograph); case 3. Incisors are absent. Premolar teeth are narrow and lack the complex apical and root structures of cheek teeth in unaffected neonatal calves. No permanent teeth have developed. (**B**) Premolar tooth; case 3. Mildly disordered ameloblasts with scant apoptosis (asterisk) adjacent to enamel (EN). Small amorphous aggregates of mineralized material, probably enamel, in papillary layer (arrow). (H&E stain). (**C**) Premolar tooth; case 3. Disorderly focus of secondary dentin or osteodentin covered with nodular aggregate of cementum projects into alveolus. (H&E stain). (**D**) Premolar tooth; case 3. New bone formation extending into alveolus. (H&E satin). (**E**) Nasal concha, case 3. Mucosal glands, which in normal calves lie between vascular cavernous stratum and the mucosal surface (asterisk), are absent. (H&E stain). (**F**) Ethmoidal concha, case 3. Olfactory glands (between arrows) are present beneath mucosal epithelium. (H&E stain). (**G**) Trachea, case 3. Note the complete absence of tracheal glands. (H&E stain).

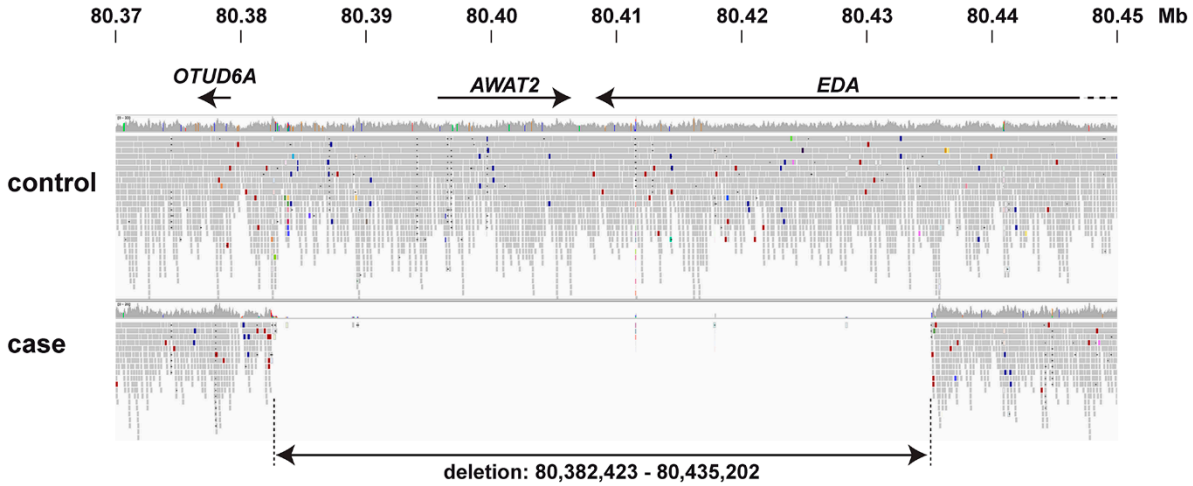


Figure 4. Genetic analysis showing screenshot from Integrative Genome Viewer (IGV) software. The short-read alignments from whole genome sequencing of an affected calf (case 3) and a normal newborn calf are shown. The affected calf lacks read coverage in central region due to a ~53 kb genomic deletion harboring the entire *AWAT2* gene and the last 6 exons of EDA.

4. Discussion

The lesions in skin, teeth and respiratory tract of affected calves closely resembled those reported in cattle with other forms of EDA-related hypohidrotic ectodermal dysplasia. Genetic analysis revealed a large X-chromosomal deletion involving the *EDA* gene.

The most striking gross abnormality in calves was in skin. Guard hairs were absent or scant over much of the trunk, neck, head and proximal part of the limbs and tail. Reports of lesions in other forms of bovine ECTD1 describe underdeveloped or hypoplastic hair follicles, fewer or abnormal sebaceous glands, absent eccrine nasolabial glands in the muzzle, and abnormal or absent apocrine sweat glands [20,22,27]. With the exception of their slightly smaller size, the predominant type of hair follicle was indistinguishable from undercoat follicles in the control calves that had normal hair coats. These may correspond to small "last formed follicles" which develop in late gestation [44]. As with ECTD1 in people and the *Tabby* mouse, the lack of functional ectodysplasin A during a key period in gestation resulted in failure of larger follicles to develop normally in most areas of the skin [17,45]. The density of hair follicles in the affected calves presented in this study was only slightly greater than in unaffected calves. This is dissimilar to a calf with ECTD1, which had five times more hair follicles than a control calf [10]—this might be due to using different skin sites (scalp vs. flank) for follicular counts in the two variants of ECTD1. Sebaceous glands in the herein studied calves were morphologically normal. Glands of the upper and lower respiratory tract were also absent, as reported in other calves with ECTD1 [9].

Apocrine sweat glands were present, generally closely associated with dermal papillae. Eccrine sweat glands are markedly reduced in ECTD1 in people, resulting in susceptibility to hyperthermia [9,35]. In previously reported cases of bovine ECTD1, apocrine glands were reported to be absent, reduced in number, slightly dilated or underdeveloped, or to have flattened epithelium suggestive of inactivity [8–10,20,22,24]. The morphology of apocrine glands in affected Red Angus-Simmental calves resembled that of control calves. The shape and size of normal bovine sweat glands is variable, determined in part by breed [46]. Unlike eccrine sweat glands of people [47] and mice [48], there is no simple way to assess the function of apocrine sweat glands in newborn calves [49]. We did not attempt to quantify sweat glands in skin from different areas.

Not surprisingly, similarities exist between the phenotype of ECTD1 in affected cattle and that of dogs, mice and humans. This includes the distribution of hairless and lightly haired areas (cattle; dog) and a predisposition to respiratory infection (cattle; dog; hu-

man) [35,50]. The distribution of hypotrichosis differs between calves and dogs in that it is generally more extensive in cattle where it generally affects facial skin, dorsal, lateral and ventral parts the neck and trunk, and the proximal portions of limbs. The distribution of ECTD1 hypotrichosis in dogs tends to be more restricted [12,37,39,51]. Although no comprehensive assessment of glands was attempted in the current study, mucosal glands were absent in upper and lower portions of the respiratory tract. As in the *Tabby* mouse [15], the olfactory glands of at least one affected Nebraska calf had developed normally. Recurrent conjunctivitis is described in some species with ECTD1 and variously ascribed to reduced tear production (dogs) [39] or a primary keratopathy (*Tabby* mouse) [33]. This has not been reported in cattle, possibly because most affected calves die or are killed so soon after birth. No clinical lesions were evident in the eyes of the five Nebraska calves. Ocular lesions were absent in one calf, which died at 10 days of age. Meibomian glands were present in his eyelids.

Deciduous incisor teeth of full-term calves erupt between birth and 21 days postpartum [43]. Incisor and canine teeth were radiographically, grossly and histologically absent in the one affected calf whose teeth were examined in detail. For these teeth, no nascent dental follicles, enamel organs, mineralized dental matrices, or dental ectomesenchyme were identified in the jaws. Premolar deciduous teeth were abnormally slender and more widely spaced than in a normal calf. The examined dental follicles of deciduous teeth did not demonstrate any overarching defect in odontoblasts, ameloblasts, cementoblasts or associated mesenchyme. This is consistent with earlier findings in Holstein calves with ECTD1 and in the *Tabby* mouse, where the principal histological dental abnormalities were oligodontia, delayed eruption, and defective dental shape, size and orientation [14,22,24]. Early in odontogenesis, however, delayed dental differentiation in *Tabby* was associated with degeneration of ameloblasts, as well as altered apoptosis patterns [52,53].

It is recognized that, in addition to abnormal development of skin and teeth, humans with ECTD1 lack normal glands in the respiratory tract, predisposing patients to nasal obstruction, sinusitis, and respiratory infections [34]. Similar changes affect the *Tabby* mouse [7] and cattle with ECTD1 [10,20,27]. Interestingly, pneumonia was responsible for the death of two of five affected calves on the affected ranch. Of the remaining three affected calves, two died without being examined postmortem to determine a cause of death. The remaining calf (case 3) was kept alive indoors until euthanasia.

The genetic investigation revealed a plausible candidate causative variant located on the X chromosome since a large part of the *EDA* gene was deleted in the sequenced affected calf. Interestingly, the *EDA* gene along with the neighboring *AWAT2* gene encoding acyl-CoA wax alcohol acyltransferase 2 was affected. *AWAT2* catalyzes the formation of wax monoesters in sebocytes [54]. The presumptive inactivation of *AWAT2* did not result in any detectable lesion in sebaceous glands. A genetic form of ECTD1 with a deletion that encompassed *EDA* and *AWAT2* has been reported in a human patient. Similar to the calves reported here, no additional clinical features were evident in that individual [55].

5. Conclusions

This represents the first form of X-linked hypohidrotic ectodermal dysplasia in cattle that affected both *EDA* and *AWAT2* genes. Regardless of the *EDA* variation, affected calves are clinically and morphologically similar, and possess many of the lesions found in affected human patients. For many genes it is known that the kind of genetic alteration influences the phenotypic outcome, e.g., the severity of a congenital defect varies or differs totally, depending on the individual variant. Interestingly for *EDA* in cattle this seems not to be the case as different kinds of variants always cause an essentially identical phenotype which is of importance for diagnostic pathologists. Furthermore, rare disorders, such as ECTD1-like syndrome in livestock are mostly not diagnosed at the molecular level, due to a lack of resources, the often-short lifespan of the animals, and concerns that producers have for their reputation as conscientious breeders. This example highlights the utility

of WGS-based precise diagnostics for understanding rare disorders in animals with an available reference genome sequence, and the need for continued surveillance for genetic disorders in cattle breeding. Genome sequencing might improve the precision of the clinicopathological diagnosis as sometimes unexpected variants in genes that were not known to be associated with a certain disorder can be detected.

Author Contributions: Conceptualization, D.O. and T.L.; methodology, D.O., C.D. and T.L.; formal analysis, D.O., C.D. and T.L.; investigation, D.O., F.L., I.M.H., C.D., B.S., L.S., B.M., C.D. and T.L.; data curation, I.M.H.; writing—original draft preparation, D.O. and T.L.; writing—review and editing, D.O., I.M.H., F.L., B.S., L.S., B.M., C.D. and T.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under awards WYO-568-16 and WYO-564-16. Furthermore, this study was partially funded by the Swiss National Science Foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but part of a clinical and pathological veterinary diagnostic case.

Data Availability Statement: The whole-genome data of the affected calf is freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA5714973.

Acknowledgments: We appreciate the cooperation of the Nebraska ranch owner for providing access to the affected herd. We acknowledge the technical assistance of Breanna Bonner, Rebecca Ashley, Rachel Griess, Jennifer McKenna, Will Laegreid and Jonathan Fox at the University of Wyoming. We acknowledge the Next Generation Sequencing Platform at the University of Bern for performing the high-throughput sequencing experiment, and the Interfaculty Bioinformatics Unit of the University of Bern for providing high performance computing infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cluzeau, C.; Hadj-Rabia, S.; Jambou, M.; Mansour, S.; Guigue, P.; Masmoudi, S.; Bal, E.; Chassaing, N.; Vincent, M.-C.; Viot, G.; et al. Only four genes (EDA1, EDAR, EDARADD, and WNT10A) account for 90% of hypohidrotic/anhidrotic ectodermal dysplasia cases. *Hum. Mutat.* **2010**, *32*, 70–72. [CrossRef] [PubMed]
2. Wright, J.T.; Fete, M.; Schneider, H.; Zinser, M.; Koster, M.I.; Clarke, A.J.; Hadj-Rabia, S.; Tadini, G.; Pagnan, N.; Visinoni, A.F.; et al. Ectodermal dysplasias: Classification and organization by phenotype, genotype and molecular pathway. *Am. J. Med. Genet. Part A* **2019**, *179*, 442–447. [CrossRef]
3. Sadier, A.; Viriot, L.; Pantalacci, S.; Laudet, V. The ectodysplasin pathway: From diseases to adaptations. *Trends Genet.* **2014**, *30*, 24–31. [CrossRef]
4. Darwin, C. *The Variation of Animals and Plants under Domestication*, 2nd ed.; John Murray: London, UK, 1875; Volume 2, p. 319.
5. Beahrs, J.O.; Lillington, G.A.; Rosan, R.C.; Russin, L.; Lindgren, J.A.; Rowley, P.T. Anhidrotic ectodermal dysplasia: Predisposition to bronchial disease. *Ann. Intern. Med.* **1971**, *74*, 92. [CrossRef] [PubMed]
6. Casal, M.L.; Jezyk, P.F.; Greek, J.M.; Goldschmidt, M.H.; Patterson, D.F. X-Linked ectodermal dysplasia in the dog. *J. Hered.* **1997**, *88*, 513–517. [CrossRef] [PubMed]
7. Del-Pozo, J.; MacIntyre, N.; Azar, A.; Headon, D.; Schneider, P.; Cheeseman, M. Role of ectodysplasin signalling in middle ear and nasal pathology in rat and mouse models of hypohidrotic ectodermal dysplasia. *Dis. Model. Mech.* **2019**, *12*, dmm037804. [CrossRef]
8. Drogemuller, C.; Distl, O.; Leeb, T. Partial deletion of the bovine ED1 gene causes anhidrotic ectodermal dysplasia in cattle. *Genome Res.* **2001**, *11*, 1699–1705. [CrossRef]
9. Drögemüller, C.; Distl, O.; Leeb, T. X-linked anhidrotic ectodermal dysplasia (ED1) in men, mice, and cattle. *Genet. Sel. Evol.* **2003**, *35*, S137–S145. [CrossRef]
10. Escouflaire, C.; Rebours, E.; Charles, M.; Orellana, S.; Cano, M.; Rivière, J.; Grohs, C.; Hayes, H.; Capitan, A. A de novo 3.8-Mb inversion affecting the EDA and XIST genes in a heterozygous female calf with generalized hypohidrotic ectodermal dysplasia. *BMC Genom.* **2019**, *20*, 1–11. [CrossRef]
11. Srivastava, A.K.; Pispa, J.; Hartung, A.J.; Du, Y.; Ezer, S.; Jenks, T.; Shimada, T.; Pekkanen, M.; Mikkola, M.L.; Ko, M.S.H.; et al. The Tabby phenotype is caused by mutation in a mouse homologue of the EDA gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13069–13074. [CrossRef]

12. Waluk, D.P.; Zur, G.; Kaufmann, R.; Welle, M.M.; Jagannathan, V.; Drögemüller, C.; Müller, E.J.; Leeb, T.; Galichet, A. A Splice Defect in the EDA Gene in dogs with an X-Linked Hypohidrotic Ectodermal Dysplasia (XLHED) Phenotype. *G3 Genes | Genomes | Genet.* **2016**, *6*, 2949–2954. [CrossRef]
13. Falconer, D.S. A Totally sex-linked gene in the house mouse. *Nat. Cell Biol.* **1952**, *169*, 664–665. [CrossRef]
14. Grüneberg, H. Genes and genotypes affecting the teeth of the mouse. *J. Embryol. Exp. Morphol.* **1965**, *14*, 137–159.
15. Grüneberg, H. The glandular aspects of the tabby syndrome in the mouse. *J. Embryol. Exp. Morphol.* **1971**, *25*, 1–19. [PubMed]
16. Casal, M.L.; Lewis, J.R.; Mauldin, E.A.; Tardivel, A.; Ingold, K.; Favre, M.; Schneider, P. Significant correction of disease after postnatal administration of recombinant ecto-dysplasin A in canine X-linked ectodermal dysplasia. *Am. J. Hum. Genet.* **2007**, *81*, 1050–1056. [CrossRef] [PubMed]
17. Gaide, O.; Schneider, P. Permanent correction of an inherited ectodermal dysplasia with recombinant EDA. *Nat. Med.* **2003**, *9*, 614–618. [CrossRef] [PubMed]
18. Margolis, C.A.; Schneider, P.; Huttner, K.; Kirby, N.; Houser, T.P.; Wildman, L.; Grove, G.L.; Schneider, H.; Casal, M.L. Prenatal treatment of X-Linked hypohidrotic ectodermal dysplasia using recombinant ectodysplasin in a canine model. *J. Pharmacol. Exp. Ther.* **2019**, *370*, 806–813. [CrossRef] [PubMed]
19. Schneider, H.; Faschingbauer, F.; Schuepbach-Mallepell, S.; Körber, I.; Wohlfart, S.; Dick, A.; Wahlbuhl, M.; Kowalczyk-Quintas, C.; Vigolo, M.; Kirby, N.; et al. Prenatal correction of X-Linked hypohidrotic ectodermal dysplasia. *N. Engl. J. Med.* **2018**, *378*, 1604–1610. [CrossRef]
20. Barlund, C.S.; Clark, E.G.; Leeb, T.; Drögemüller, C.; Palmer, C.W. Congenital hypotrichosis and partial anodontia in a crossbred beef calf. *Can. Vet. J.* **2007**, *48*, 612–614. [PubMed]
21. Braun, U.; Ansari, A.H.; Hediger, R.; Süss, U.; Ehrensperger, F. Hypotrichosis and oligodontia, combined with an Xq-deletion, in a calf of the Swiss Holstein breed. *Tierarztl. Prax.* **1988**, *16*, 39–44. [PubMed]
22. Drögemüller, C.; Peters, M.; Pohlenz, J.; Distl, O.; Leeb, T. A single point mutation within the ED1 gene disrupts correct splicing at two different splice sites and leads to anhidrotic ectodermal dysplasia in cattle. *J. Mol. Med.* **2002**, *80*, 319–323. [CrossRef]
23. Gargani, M.; Valentini, A.; Pariset, L. A novel point mutation within the EDA gene causes an exon dropping in mature RNA in Holstein Friesian cattle breed affected by X-linked anhidrotic ectodermal dysplasia. *BMC Vet. Res.* **2011**, *7*, 35. [CrossRef]
24. Karlskov-Mortensen, P.; Cirera, S.; Nielsen, O.L.; Arnbjerg, J.; Reibel, J.; Fredholm, M.; Agerholm, J.S. Exonization of a LINE1 fragment implicated in X-linked hypohidrotic ectodermal dysplasia in cattle. *Anim. Genet.* **2011**, *42*, 578–584. [CrossRef]
25. Ogino, A.; Kohama, N.; Ishikawa, S.; Tomita, K.; Nonaka, S.; Shimizu, K.; Tanabe, Y.; Okawa, H.; Morita, M. A novel mutation of the bovine EDA gene associated with anhidrotic ectodermal dysplasia in Holstein cattle. *Hereditas* **2011**, *148*, 46–49. [CrossRef] [PubMed]
26. Ogino, A.; Shimizu, K.; Tanabe, Y.; Morita, M. De novo mutation of the bovine EDA gene associated with anhidrotic ectodermal dysplasia in Japanese Black cattle. *Anim. Genet.* **2011**, *43*, 646. [CrossRef] [PubMed]
27. Seeliger, F.; Drögemüller, C.; Tegtmeier, P.; Baumgartner, W.; Distl, O.; Leeb, T. Ectodysplasin-1 Deficiency in a german holstein bull associated with loss of respiratory mucous glands and chronic rhinotracheitis. *J. Comp. Pathol.* **2005**, *132*, 346–349. [CrossRef]
28. Bourneuf, E.; Otz, P.; Pausch, H.; Jagannathan, V.; Michot, P.; Grohs, C.; Piton, G.; Ammermüller, S.; Deloche, M.C.; Fritz, S.; et al. Rapid discovery of de novo deleterious mutations in cattle enhances the value of livestock as model species. *Sci. Rep.* **2017**, *7*, 1–19. [CrossRef]
29. Cole, L.J. A Defect of hair and teeth in cattle—Probably hereditary1. *J. Hered.* **1919**, *10*, 303–306. [CrossRef]
30. Denis, B.; Théret, M.; Blin, P.; Bernard, C.; Lauvergne, J. Hypotrichose congénitale en race bovine normande. I.—Étude descriptive. *Ann. Génétique Sélection Anim.* **1975**, *7*, 251–261. [CrossRef] [PubMed]
31. Drieux, H.; Priouzeau, M.L.; Thiery, G. Congenital hypotrichosis with absence of teeth and horns, and macroglossia in a calf. *Recl. Med. Vet.* **1950**, *126*, 385–399.
32. Wijeratne, W.V.; O'Toole, D.; Wood, L.; Harkness, J.W. A genetic, pathological and virological study of congenital hypotrichosis and incisor anodontia in cattle. *Vet. Rec.* **1988**, *122*, 149–152. [CrossRef]
33. Cui, C.-Y.; Smith, J.A.; Schlessinger, D.; Chan, C.-C. X-Linked anhidrotic ectodermal dysplasia disruption yields a mouse model for ocular surface disease and resultant blindness. *Am. J. Pathol.* **2005**, *167*, 89–95. [CrossRef]
34. Mehta, U.; Brunworth, J.; Fete, T.J.; Sindwani, R. Head and neck manifestations and quality of life of patients with ectodermal dysplasia. *Otolaryngol. Head Neck Surg.* **2007**, *136*, 843–847. [CrossRef]
35. Wohlfart, S.; Meiller, R.; Hammersen, J.; Park, J.; Menzel-Severing, J.; Melichar, V.O.; Huttner, K.; Johnson, R.; Porte, F.; Schneider, H. Natural history of X-linked hypohidrotic ectodermal dysplasia: A 5-year follow-up study. *Orphanet J. Rare Dis.* **2020**, *15*, 1–11. [CrossRef]
36. Anbouba, G.M.; Carmany, E.P.; Natoli, J.L. The characterization of hypodontia, hypohidrosis, and hypotrichosis associated with X-linked hypohidrotic ectodermal dysplasia: A systematic review. *Am. J. Med. Genet. Part A* **2020**, *182*, 831–841. [CrossRef]
37. Moura, E.; Cirio, S.M. Clinical and genetic aspects of X-linked ectodermal dysplasia in the dog—A review including three new spontaneous cases. *Vet. Dermatol.* **2004**, *15*, 269–277. [CrossRef] [PubMed]
38. Lewis, J.R.; Reiter, A.M.; Mauldin, E.A.; Casal, M.L. Dental abnormalities associated with X-linked hypohidrotic ectodermal dysplasia in dogs. *Orthod. Craniofacial Res.* **2010**, *13*, 40–47. [CrossRef]
39. Moura, E.; Rotenberg, I.S.; Pimpão, C.T. X-Linked hypohidrotic ectodermal dysplasia—General features and dental abnormalities in affected dogs compared with human dental abnormalities. *Top. Companion Anim. Med.* **2019**, *35*, 11–17. [CrossRef] [PubMed]

40. Cornish, T.E.; Van Olphen, A.L.; Cavender, J.L.; Edwards, J.M.; Jaeger, P.T.; Vieyra, L.L.; Woodard, L.F.; Miller, D.R.; O'Toole, D. Comparison of ear notch immunohistochemistry, ear notch antigen-capture ELISA, and buffy coat virus isolation for detection of calves persistently infected with bovine viral diarrhea virus. *J. Vet. Diagn. Investig.* **2005**, *17*, 110–117. [CrossRef]
41. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
42. Robinson, J.T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E.S.; Getz, G.; Mesirov, J.P. Integrative genomics viewer. *Nat. Biotechnol.* **2011**, *29*, 24–26. [CrossRef]
43. Brown, W.A.; Christofferson, P.V.; Massler, M.; Weiss, M.B. Postnatal tooth development in cattle. *Am. J. Vet. Res.* **1960**, *21*, 7–34.
44. Lyne, A.; Heideman, M.J. The Pre-Natal Development of Skin and hair in Cattle (*Bos Taurus* L.). *Aust. J. Biol. Sci.* **1959**, *12*, 72–95. [CrossRef]
45. Cui, C.Y.; Kunisada, M.; Childress, V.; Michel, M.; Schlessinger, D. Shh is required for Tabby hair follicle development. *Cell Cycle.* **2011**, *10*, 3379–3386. [CrossRef]
46. McEwan Jenkinson, D.; Nay, T. The sweat glands and hair follicles of European cattle. *Aust. J. Biol. Sci.* **1972**, *25*, 585–595. Available online: https://www.publish.csiro.au/bi/pdf/bi9720585 (accessed on 3 January 2021). [CrossRef] [PubMed]
47. Rouse, C.; Siegfried, E.; Breer, W.; Nahass, G. Hair and sweat glands in families with hypohidrotic ectodermal dysplasia: Further characterization. *Arch. Dermatol.* **2004**, *140*, 850–855. [CrossRef] [PubMed]
48. Blecher, S.R. Anhidrosis and absence of sweat glands in mice hemizygous for the tabby gene; supportive evidence for the hypothesis of homology between tabby and human anhidrotic (hypohidrotic) ectodermal dysplasia (Christ-Siemens-Touraine Syndrome). *J. Investig. Dermatol.* **1986**, *87*, 720–722. [CrossRef] [PubMed]
49. Pan, Y.; Donegan, S.; Hayman, R. Sweating rate at different body regions in cattle and its correlation with some quantitative components of sweat gland volume for a given area of skin. *Aust. J. Agric. Res.* **1969**, *20*, 395–403. [CrossRef]
50. Casal, M.L.; Mauldin, E.A.; Ryan, S.; Scheidt, J.L.; Kennedy, J.; Moore, P.F.; Felsburg, P.J. Frequent respiratory tract infections in the canine model of X-linked ectodermal dysplasia are not caused by an immune deficiency. *Vet. Immunol. Immunopathol.* **2005**, *107*, 95–104. [CrossRef] [PubMed]
51. Vasiliadis, D.; Hewicker-Trautwein, M.; Klotz, D.; Fehr, M.; Ruseva, S.; Arndt, J.; Metzger, J.; Distl, O. A de novo EDA-Variant in a litter of shorthaired standard dachshunds with X-Linked hypohidrotic ectodermal dysplasia. *G3 Genes | Genomes | Genet.* **2019**, *9*, 95–104. [CrossRef]
52. Boran, T.; Lesot, H.; Peterka, M.; Peterkova, R. Increased apoptosis during morphogenesis of the lower cheek teeth in Tabby/EDA Mice. *J. Dent. Res.* **2005**, *84*, 228–233. [CrossRef] [PubMed]
53. Sofaer, J.A. Aspects of the tabby-crinkled-downless syndrome. I. The development of tabby teeth. *J. Embryol. Exp. Morphol.* **1969**, *22*, 181–205.
54. Yen, C.L.E.; Brown, C.H., IV; Monetti, M.; Farese, R.V., Jr. A human skin multifunctional O-acyltransferase that catalyzes the synthesis of acylglycerols, waxes, and retinyl esters. *J. Lipid Res.* **2005**, *46*, 2388–2397. [CrossRef] [PubMed]
55. Chaudhary, A.K.; Sankar, V.; Bashyam, M.D.; Bashyam, D. A novel large deletion that encompasses EDA and the downstream gene AWAT2 causes X-linked hypohidrotic/anhidrotic ectodermal dysplasia. *J. Dermatol. Sci.* **2016**, *84*, 105–107. [CrossRef] [PubMed]

A heterozygous missense variant in the *COL5A2* in Holstein cattle resembling the classical Ehlers–Danlos syndrome

Journal:	Animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, co-writing draft manuscript, writing – review and editing
Displayed version:	published version
DOI:	10.3390/ani10112002

animals

MDPI

Article

A Heterozygous Missense Variant in the *COL5A2* in Holstein Cattle Resembling the Classical Ehlers–Danlos Syndrome

Joana G. P. Jacinto [1,2,†], Irene M. Häfliger [2,†], Inês M. B. Veiga [3,†], Anna Letko [2], Cinzia Benazzi [1], Marilena Bolcato [1] and Cord Drögemüller [2,*]

1 Department of Veterinary Medical Sciences, University of Bologna, 40064 Ozzano Emilia (Bologna), Italy; joana.goncalves2@studio.unibo.it (J.G.P.J.); cinzia.benazzi@unibo.it (C.B.); marilena.bolcato2@unibo.it (M.B.)
2 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch (I.M.H.); anna.letko@vetsuisse.unibe.ch (A.L.)
3 Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; ines.veiga@vetsuisse.unibe.ch
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch
† These authors contributed equally to the work.

Received: 13 October 2020; Accepted: 28 October 2020; Published: 30 October 2020

Simple Summary: Genodermatoses represent inherited disorders of the skin that mostly follow a monogenic mode of inheritance. Heritable connective tissue disorders such as classical Ehlers–Danlos syndrome (cEDS) belong to this group of human rare diseases that sporadically occur in other species. Herein, affected cattle are reported showing skin lesions including cutis laxa clinically and pathologically resembling cEDS in humans. Microscopic findings in the deeper dermis were consistent with collagen dysplasia. Whole-genome sequencing (WGS) identified a most likely disease-causing mutation in the *COL5A2* gene. The *COL5A2* gene is known to be associated with dominant inherited cEDS forms in mice and humans, but so far, it was not shown to cause a similar phenotype in domestic animals. The disease phenotype examined herein showed co-segregation with the identified missense variant within the maternal line across two generations and is most likely due to a spontaneous mutation event. Rare non-lethal disorders such as cEDS in livestock are mostly not diagnosed, but might affect animal welfare and thus lower the value of affected animals. WGS-based precision diagnostics allows understanding rare disorders and supports the value of surveillance of cattle breeding populations for harmful genetic disorders.

Abstract: Classical Ehlers–Danlos syndrome (cEDS) is a heritable connective tissue disorder characterized by variable degrees of skin hyperextensibility and fragility, atrophic scarring, and generalized joint hypermobility. The purpose of this study was to characterize the clinicopathological phenotype of a cEDS-affected Holstein calf and to identify the causative genetic variant associated with the disorder by whole-genome sequencing (WGS). A 3-day-old female Holstein calf was referred because of easily induced skin detachment and hyperextensibility in the neck. A complete clinical investigation was performed in the calf, dam, and maternal-grandmother. The calf and dam showed hyperextensibility of the neck skin and atrophic scarring; additionally, the calf presented skin fragility. Moreover, the histopathology of biopsies from the calf and its dam showed that the collagen bundles in affected skin areas were wavy, short, thin, and surrounded by edema and moderate to severe acute hemorrhages. Genetic analysis revealed a private heterozygous missense variant in *COL5A2* (c.2366G>T; p.Gly789Val) that was present only in the calf and dam. This confirmed the diagnosis of cEDS and represents the first report of a causal variant for cEDS in cattle and the first *COL5A2*-related large animal model.

Keywords: *Bos taurus*; collagen dysplasia; collagen V; connective tissue; precision medicine; skin fragility; whole-genome sequencing

1. Introduction

Sporadically occurring, genodermatoses represent inherited disorders of the skin that mostly follow a monogenic mode of inheritance in livestock animals such as cattle [1]. Heritable connective tissue disorders, e.g., Ehlers–Danlos syndrome (EDS), belong to this group of human rare diseases. EDS encompasses a clinically- and heritably-heterogeneous group of connective tissue disorders (Online Mendelian Inheritance in Man (OMIM PS130000)) (https://www.omim.org/phenotypicSeries/PS130000) characterized by a variable degree of skin hyperextensibility, joint hypermobility, and tissue fragility. Currently, human EDS classification distinguishes 13 subtypes and 19 different associated genes mainly involved in collagen and extracellular matrix synthesis and maintenance reflecting the clinical and genetic heterogeneity [2]. Human EDS forms are grouped based on the underlying pathogenetic mechanisms related to primary structure and processing of collagen (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A1*, *COL5A2*, *ADAMTS2*), collagen folding and cross-linking (*PLOD1*, *FKBP14*), structure and function of the myomatrix (*TNXB*, *COL12A1*), glycosaminoglycan biosynthesis (*B4GALT7*, *B3GALT6*, *CHST14*, *DSE*), complement pathway (*C1S*, *C1R*), and intracellular processes (*SLC39A13*, *ZNF469*, *PRDM5*) [3]. Classical EDS (cEDS) in humans is a rare autosomal dominant disorder predominantly associated with a deficiency of type V collagen (COLLV) encoded by the *COL5A1* and *COL5A2* genes, which is a quantitatively minor fibrillar collagen that presents a nearly ubiquitous distribution in a variety of connective tissues [4].

Various forms of EDS have been identified in many animal species (OMIA000327), including horses [5,6], dogs [7–9], cats [10], mink [11], rabbits [12], sheep [13–15], and cattle (OMIA000328-9913 (https://www.omia.org/OMIA000328/9913/); OMIA001716-9913 (https://www.omia.org/OMIA001716/9913/)) [16–18]. Pathogenic variants causing forms of EDS in animals have been identified in known candidate genes for EDS (*COL5A1*, *ADAMTS2*, *PLOD1*) [5,7,8,10,13,15,18], or novel genes (*EPYC*, *TNBX*, *PPIB)* discovered in EDS-affected domestic animals [6,9,17]. This highlights the potential of studying inherited conditions in such species to assign a role or function to previously uncharacterized genes or to add additional functions to known genes in regard to skin development [1].

In this study, we aimed to characterize the clinical and pathological phenotypes of a cEDS-affected Holstein calf and its dam, and to identify the causative genetic variant associated with the disorder using whole-genome sequencing (WGS).

2. Materials and Methods

2.1. Ethics Statement

This study did not require official or institutional ethical approval as it was not experimental, but rather part of clinical and pathological veterinary diagnostics. All animals in this study were examined with the consent of their owners and handled according to good ethical standards.

2.2. Clinicopathological Investigation

A 3-day-old female Holstein calf was referred by the farm veterinarian because of easily induced skin detachment in the neck and skin hyperextensibility shortly after birth. Upon specific request, the owner informed that its 3-year-old dam and its 6-year-old maternal-grandmother had also been presenting skin alterations for a long time, but was not able to specify since when. All three animals, the calf, its dam, and its maternal-grandmother, were clinically examined and a complete blood count (CBC) and blood chemistry profile were obtained. Blood samples from the calf and its dam were sent for routine viral and parasitological analysis (bovine viral diarrhea, bovine Schmallenberg virus,

bluetongue virus, *Neospora* spp., *Toxoplasma* spp.) using antigen-enzyme-linked immunosorbent assay (ELISA) and antibody-polymerase chain reaction (PCR). Three weeks later (calf's age = 29 days), skin biopsies using an 8 mm punch were obtained from the ulcerated cervical skin, from the skin surrounding the cervical ulceration, and from normal skin from the neck from the calf, as well as from the altered cervical skin from its dam. The collected samples were fixed in 10% neutral buffered formalin, trimmed, processed, embedded in paraffin wax, sectioned at 4 µm, and stained with haematoxylin and eosin (H&E) for further histological evaluation. Further clinical control was carried out after six months (calf's age = 7 months). All animals were housed in a freestall system.

2.3. DNA Samples

Genomic DNA was isolated from ethylenediaminetetraacetic acid (EDTA) blood samples from the calf and its dam; from EDTA blood sample, normal skin, and lesioned skin of the maternal-grandmother; and from the semen of the sire using Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland).

2.4. Whole-Genome Sequencing

WGS using the Illumina NovaSeq6000 (Illumina Inc., San Diego, CA, USA) was performed on the genomic DNA of the calf. The sequenced reads were mapped to the ARS-UCD1.2 reference genome, resulting in an average read depth of approximately 17× [19], and single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format files were in accordance with the 1000 Bull Genomes Project processing guidelines of run 7 (www.1000bullgenomes.com) [20], except for the trimming, which was performed using fastp [21]. Further preparation of the genomic data was done according to Häfliger et al., 2020 [22]. In order to find private variants, we compared the genotypes of the affected calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that are publicly available (Table S1) in the European Nucleotide Archive (SAMEA7015115 is the sample accession number of the affected calf; http://www.ebi.ac.uk/en). Integrative Genomics Viewer (IGV) [23] software was used for visual inspection of genome regions containing possible candidate genes.

2.5. Targeted Genotyping

Polymerase chain reaction (PCR) and Sanger sequencing were used to validate and genotype the variant identified from WGS. PCR products from genomic DNA were amplified using AmpliTaq Gold 360 Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) and the PCR amplicons were directly sequenced on an ABI3730 capillary sequencer (Thermo Fisher Scientific, Darmstadt, Germany). The *COL5A2* missense variant (XM_024979774.1:g.7331916G>T) was genotyped using the following primers: 5′- ACCAGGGCTTCAAGGTATGC-3′ (forward primer) and 5′-CACCATGGGAACATGAGGCT-3′ (reverse primer). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA).

2.6. Protein Predictions

PROVEAN [24], MutPred2 [25], and PredictSNP1 [26] were used to predict the biological consequences of the discovered variant on protein. For multispecies sequence alignments, the following National Center for Biotechnology Information (NCBI) proteins accessions were used: XP_024835542.1 (*Bos taurus*), NP_000384.2 (*Homo sapiens*), XP_001164152.1 (*Pan troglodytes*), XP_002799008.1 *(Macaca mulatta*), XP_005640450.1 (*Canis lupus*), NP_031763.2 (*Mus musculus*), NP_445940.1 (*Rattus norvegicus*), XP_004942453.1 (*Gallus gallus*), NP_001139254.1 (*Danio rerio*), and XP_002931546.2 (*Xenopus tropicalis*).

2.7. Sequence Accessions

All references to the bovine *COL5A2* gene correspond to the NCBI accessions NC_037329.1 (chromosome 2, ARS-UCD1.2), XM_024979774 (*COL5A2* mRNA), and XP_024835542.1 (COL5A2 protein). For the protein structure of COL5A2, the Uniprot database (https://www.uniprot.org/) accession number A0A3Q1MDT9 was used.

3. Results

3.1. Clinical Phenotype

On clinical investigation, the calf, its dam, and its maternal-grandmother were found to be clinically healthy with the exception of the skin alterations. Particular clinical examination of the cardiovascular, respiratory, urinary, musculoskeletal, and nervous systems showed no abnormalities. Moreover, no joint hypermobility was observed. The blood investigation of the calf showed a moderate monocytosis (1760/mm^3), a mild neutrophilia (4650/mm^3), a mild hypocholesterolemia (47 mg/dL), and a mild hypoproteinemia (6.10 g/dL) with hypoalbuminemia (2.82 g/dL). No abnormalities were detected in the CBC and chemistry profiles of the dam and maternal-grandmother. Blood viral analysis revealed positivity for bovine Schmallenberg virus using antigen-ELISA in the calf, and positivity for bovine viral diarrhea using antigen-ELISA in the calf and its dam. The animals tested negative for all the remaining viral and parasitological analyses.

The integumentary system examination at 3 days of age of the calf revealed a symmetrical, bilateral ulceration secondary to minor trauma, delimitated cranially by the occipital region, caudally by the cranial margin of the scapula, and ventrally by the sternocephalic muscles (Figure 1a). The ulcerated surface was dry, non-painful, non-bleeding, and pinkish with the presence of purulent material at the edges. During palpation, at the wound edges, a spontaneous detachment of the skin from the subcutaneous tissue was noticed. Furthermore, the animal presented hyperextensibility of the skin mostly in the neck. At 29 days of age, the previously observed neck wound was dry, scabby, and crustose with the edges of the lesion firmly embedded in the subcutaneous tissue; atrophic scarring was also present (Figure 1b). At seven months, the calf's lesions were similar to its dam, and characterized by multiple wrinkles, folds, papyraceous scars, cutis laxa, and hyperextensibility of the neck skin (Figure 1c).

Examination of the dam revealed multiple wrinkles, folds, papyraceous scars, cutis laxa, and hyperextensibility of the neck skin (Figure 2a,b). Moreover, its maternal-grandmother showed milder skin lesions in the ventral part of the neck at the level of the larynx, characterized by the presence of bald areas and scabs (Figure S1).

Based on these clinical observations, the calf and its dam were consequently suspected to suffer from cEDS, while the skin lesions observed in the maternal-grandmother were suspected to be acquired as the skin has been traumatised by recurring long-term mechanical stress of the feed fence.

3.2. Histopathological Phenotype

Histologically, the epidermis at the level of the cervical lesion of the calf displayed a severe, diffuse ulceration, with underlying proliferation of mature granulation tissue associated with neovascularisation, abundant neutrophilic superficial infiltrates, and adnexal structure loss (Figure 3a). The epidermis at the border of the ulcerated area displayed abundant serocellular crusts associated with serum lake formation, spongiosis, and ballooning degeneration of the stratum granulosum. Multifocal perivascular, moderate lymphoplasmacytic infiltrates could be observed in the dermis underlying the ulcerated area. The deeper dermis displayed interlacing, wavy, short, thin collagen bundles, which were surrounded by a moderate interstitial edema, as well as moderate to severe acute hemorrhages (Figure 3a).

Figure 1. Neck skin lesions of the classical Ehlers–Danlos syndrome (cEDS)-affected Holstein calf. (**a**) Calf age = 3 days: severe, extensive ulceration secondary to minor trauma; note the spontaneous detachment of the skin from the subcutaneous tissue. (**b**) Calf age = 29 days: atrophic scarring; note the scabby and crustose wound with the edges firmly embedded in the subcutaneous tissue. (**c**) Calf age = 7 months: papyraceous scars and cutis laxa.

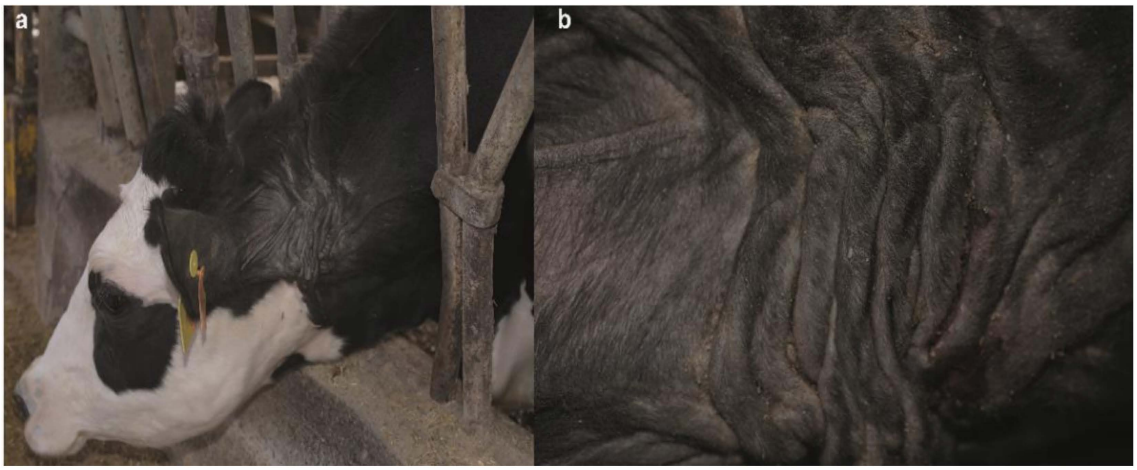


Figure 2. Neck skin lesions of the cEDS-affected Holstein dam. (**a**) Note the multiple wrinkles, folds, papyraceous scars and cutis laxa. (**b**) Details of the neck skin lesion.

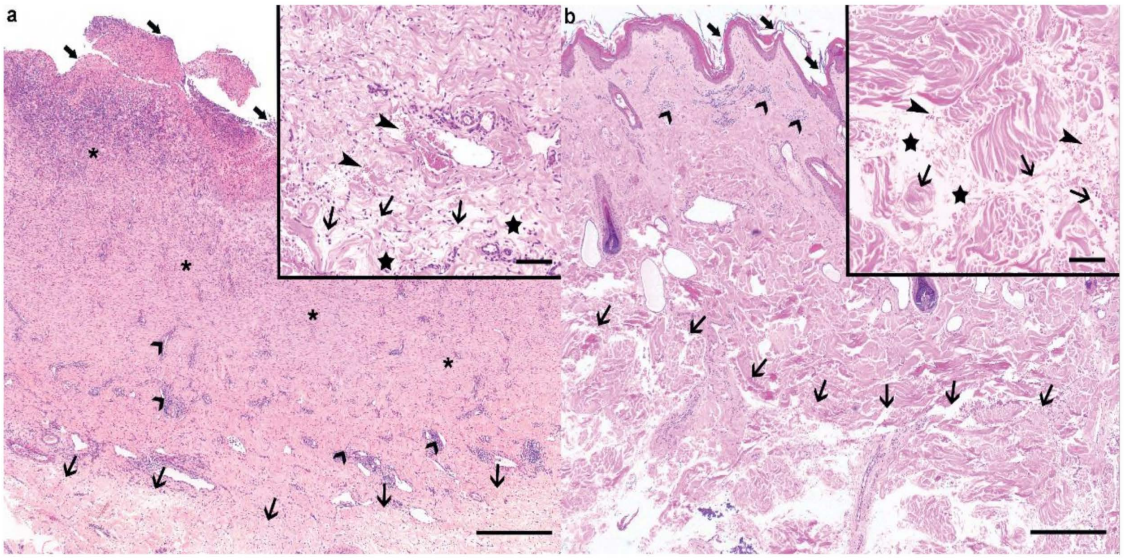


Figure 3. Histological findings of the cEDS-affected Holstein calf and its dam. (**a**) The ulcerated (large arrows) cervical lesion from the calf displayed a prominent granulation tissue proliferation (asterisks) associated with neovascularization and severe neutrophilic infiltration. Moderate, lymphoplasmacytic perivascular infiltrates were visible in the superficial dermis (large arrowheads). Within the deeper dermis, the collagen bundles were loose and irregular (thin arrows). Haematoxylin and eosin (H&E), bar 500 µm. Inset: higher magnification of the affected connective tissue within the deeper dermis. The collagen bundles were wavy, short, and thin (thin arrows), and surrounded by edema (stars) and acute hemorrhage (thin arrowheads) in the absence of vascular changes. Haematoxylin and eosin (H&E), bar 100 µm. (**b**) In the dam, the epidermis was irregular, mildy hyperplastic, and layered by large amount of lamellar to compact, orthokeratotic keratin (large arrows). Mild to moderate, perivascular lymphocytic and plasmacellular infiltrates were visible in the superficial dermis (large arrowheads), while similar changes to the ones described in the calf could be observed in the deeper dermis (thin arrows). Haematoxylin and eosin (H&E), bar 500 µm. Inset: Higher magnification of the affected connective tissue within the deeper dermis. Similar changes to the ones observed in the calf could be observed, namely, wavy, short, and thin collagen bundles (thin arrows), interstitial edema (stars), and acute hemorrhage (thin arrowheads). Haematoxylin and eosin (H&E), bar 100 µm.

In the punch biopsy taken from the normal skin from the neck, the epidermis displayed a normal thickness and was covered by a large amount of fairly compact, orthokeratotic keratin. Mild to severe, interstitial eosinophilic and neutrophilic infiltrates of unknown origin were present mostly within the deeper dermis and did not allow the identification of the dermal changes observed at the site of the ulceration.

The cutaneous punch biopsy taken from the neck of the dam displayed a mildly hyperplastic and irregular epidermis, and was covered by a large amount of lamellar to compact, orthokeratotic keratin (Figure 3b). The superficial dermis displayed mild to moderate perivascular infiltrates composed of lymphocytes, plasma cells, and occasionally eosinophils. The deeper dermis displayed similar changes to the ones observed in the biopsies from the calf (Figure 3b).

The histological findings in the deeper dermis from both the calf and its dam were compatible with collagen dysplasia within the deeper dermis, and thus with the clinical suspicion of cEDS.

3.3. Genetic Analysis

Filtering of WGS for private variants present in the affected calf and absent in the 496 available control genomes identified a single protein-changing variant in COL5A2 with a predicted moderate impact on the encoded protein. The heterozygous variant in COL5A2 exon 35 on chromosome 2 (chr2:g.7331916G>T) was confirmed using IGV software (Figure 3a,c). The detected COL5A2 variant

(XM_024979774.1: c.2366G>T) alters the encoded amino acid of COL5A2 residue 789 of the collagen triple-helical region (XP_024835542.1: p.Gly789Val) included in the collagen alpha-2(V) chain (Figure 4e). Furthermore, the glycine to valine substitution affects an evolutionary conserved residue (Figure 4f) and was predicted to be deleterious (PROVEAN score: −8.294; MutPred2 score: 0.917; PredictSNP1 score: 0.869). To confirm and evaluate the presence of the COL5A2 variant, the affected genomic region was amplified by PCR and Sanger sequenced in the calf, its dam, its maternal-grandmother, and its sire (Figure 4b). Analyzing the sequencing data, we observed that the calf and its dam were heterozygous for the detected COL5A2 variant, whereas the sire and the maternal-grandmother were homozygous for the wild type allele (Figure 4b,d). Unfortunatly, no samples from other closely related animals such as the maternal-grandfather were available.

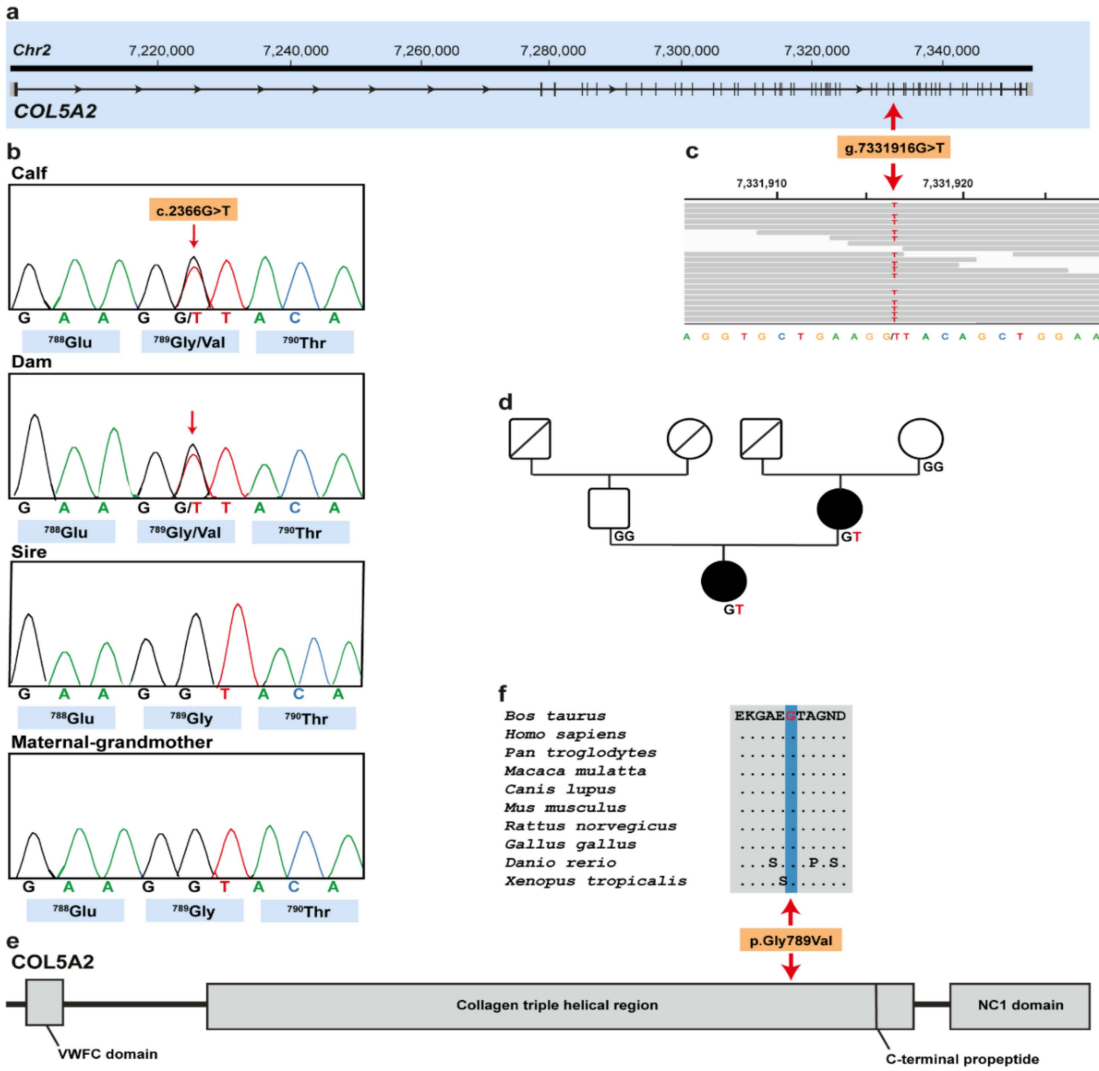


Figure 4. cEDS-associated *COL5A2* missense variant in Holstein cattle. (**a**) *COL5A2* gene structure showing the variant location on chromosome 2, exon 35 (red arrow). (**b**) Electropherograms of the calf, dam, sire, and maternal-grandmother. (**c**) Integrative Genomics Viewer (IGV) screenshot presenting the g.7331916G>T variant in the affected calf. (**d**) Pedigree of the cEDS-affected Holstein family. Males are represented by squares and females by circles. Affected animals are represented with full black symbols, while non-affected animals are represented by full white symbols. Unknown genotypes are represented by symbols with a diagonal line. *COL5A2* genotypes are shown for all available animals. (**e**) Schematic representation of COL5A2 protein and its three functional domains. (**f**) Multiple sequence alignment of the collagen alpha-2(V) chain of the COL5A2 protein encompassing the region of the p.Gly789Val variant demonstrates complete evolutionary conservation across species.

4. Discussion

The identified missense *COL5A2* variant affects a functionally important site of an obvious candidate gene and thus represents the most likely pathogenic variant associated with the observed cEDS phenotype of two examined Holstein cattle family members. In veterinary medicine, so far, two distinct pathogenic variants in *COL5A1* associated with cEDS have been reported in dogs [8] and cats [10]. To the best of our knowledge, no pathogenic variant in the *COL5A2* associated with cEDS has been reported in domestic animal species. Therefore, this study in cattle provides the first example of a *COL5A2*-related congenital skin disorder in domestic animals.

In human medicine, mutations in *COL5A2* are associated with autosomal dominantly inherited cEDS type 2 (OMIM 130010) (https://omim.org/entry/130010). The most recent update of the Leiden Open Variation Database (LOVD) lists 312 different pathogenic variants that affect COLLV [27]. In particular, 220 distinct *COL5A1* and 92 *COL5A2* pathogenic variants are described [28]. Diagnosis of cEDS in humans relies on fulfilling minimal criteria that encompass skin hyperextensibility and atrophic scaring plus either another major criterion, generalized joint hypermobility, and/or at least three minor criteria (Table 1), as well as mandatory molecular test confirmation [2].

Table 1. Classification of human classical Ehlers–Danlos syndrome (cEDS).

Inheritance	Autosomal dominant
Molecular Basis	*COL5A1; COL5A2; COL1A1; COL3A1*
Major Criteria	1. Skin hyperextensibility and atrophic scarring 2. Generalized joint hypermobility
Minor Criteria	1. Easy bruising 2. Soft, doughy skin 3. Skin fragility (or traumatic splitting) 4. Molluscoid pseudotumors 5. Subcutaneous spheroids 6. Hernia 7. Epicanthal folds 8. Complications of joint hypermobility 9. Family history of first-degree relative who meets clinical criteria

Adapted from Malfait et al. 2017 [2].

As described before, the calf and its dam carry a deleterious heterozygous missense variant in the *COL5A2* gene. More than 90% of cEDS human patients harbor a heterozygous mutation in one of the genes encoding COLLV [4,29,30], such as, for example, the missense variant p.Gly934Arg in *COL5A2*, where there is a substituition of a glycine residue within the triple helical domain (Gly-X-Y) [31].

The clinical and pathological phenotype in both cases of this study resembled a form of cEDS. The affected Holstein calf and its dam met one major criterion (skin hyperextensibility and atrophic scarring) of the human classification system (Table 1). Furthermore, one minor criterion was met in the calf (skin fragility/traumatic splitting). Even though the two presented cases do not completely fulfill the clinical criteria for human cEDS classification, the diagnosis of cEDS has been assumed. It is worth noting that, in veterinary medicine, just a few cases have been reported in the literature, rendering it difficult to develop such a classification system adapted for domestic animal species. In addition, the cases presented herein show lesions restricted to the neck region. Therefore, this may be a specific characteristic of bovine cEDS. Furthermore, it is assumed that the skin lesions observed in the maternal-grandmother were linked to trauma. In fact, the prevalalence of neck skin lesions related to the infrastructure of freestall farms is around 9% [32]. Environmental factors cause phenocopies, which are incidents in which non-genetic conditions simulate a genetic disorder. On the farm where the three animals were housed, there were more cows showing neck skin lesions similar to the maternal-grandmother's that may represent phenocopies due to recurring long-term mechanical stress of the feed fence. However, the calf's owner did not recall having any similarly affected animals in the past as the cEDS-affected calf and its dam.

We speculate that the mutation either occurred post-zygotically during the fetal development of the affected dam or represents a germline mutation that occurred in the maternal-grandfather. Nonetheless, this *de novo* mutation was then transmitted to the cEDS-affected calf. The amplification of the mutated allele in the maternal-grandmother using DNA extracted from EDTA blood, skin from the neck lesions area, and normal skin resulted in homozygous wild-type status. Therefore, the maternal-grandmother has been excluded as a mosaic ancestor. However, to prove that the identified mutation in *COL5A2* indeed occured *de novo*, genotyping of the maternal-grandsire would be needed.

Moreover, the identified deleterious variant and the conservation of the affected glycine amino acid residue of COL5A2 at position 789 in the highly conserved triple-helical domain also suggest that this variant is most likely pathogenic. The predicted amino acid exchange occurs in the protein triple-helical domain, and by analogy with helix glycine substitutions; for example, in collagen alpha-1(I) in osteogenesis imperfecta (OMIM 120150) (https://www.omim.org/entry/120150), in collagen alpha-2(II) in chondrodysplasias (OMIM 120160) (https://www.omim.org/entry/120160), and in collagen alpha-3(III) and -4(IV) in Alport syndrome II (OMIM 203780) (https://www.omim.org/entry/203780), it would be expected to disrupt the propagation of collagen triple helix, resulting in abnormal molecules linked to the disorder [33]. Collagen triple helix folding and stability is critically dependent on having a glycine as every third amino acid in the triplet repeat sequence (Gly-x-y). Therefore, replacing glycine with a bulky amino acid (in this case, a valine) has the potential to disrupt helix folding and lead to increased posttranslational lysine hydroxylation and glycosylation, compromising the triple helix structural integrity and retention of the mutant trimers in the endoplasmic reticulum, which can have an impact in the cellular function [34]. In addition, cellular quality control mechanisms that bring about endoplasmic reticulum-retention and degradation of misfolded collagens are leaky and collagen hetero- or homotrimers containing one or several mutant pro-alpha-chains are often secreted, having an important predicted impact on collagen fibril formation and stability, and altered interactions with other extracellular matrix components [34]. The major variant of COLLV is a heterotrimer composed of two pro-alpha-1 chains and a single pro-alpha-2 chain, which are encoded by the *COL5A1* and *COL5A2* genes, respectively [35]. COLLV plays a central role in the assembly of tissue-specific matrices. Several COLLV isoforms have been reported, however, the most widely accepted form is the [α1(V)]2α2(V) heterotrimer that co-assembles with type I collagen into heterotypic type I/V collagen fibrils in the extracellular matrix. COLLV is thought to regulate the diameter of these fibrils by retention of its large N-propeptide domain, which projects above the surface of the collagen fibril [35].

5. Conclusions

Rare non-lethal disorders such as cEDS in livestock are usually not reported or diagnosed when the animals show mild to moderate phenotype, but they affect animal welfare through secondary wounds and thus lower the value of the affected animals. Additionally, molecular diagnosis is often not performed because of a lack of resources and diagnostic tools, and/or low value of the animals.

Investigation of these cases allowed a complete clinical, pathological, and molecular genetic study, enabling for the first time the diagnosis of a dominantly inherited cEDS form in a family of Holstein cattle associated with a *COL5A2* variant. Furthermore, this example highlights the utility of WGS-based precision diagnostics for understanding rare disorders in animals with an available reference genome sequence and the value of surveillance of cattle breeding populations for harmful genetic disorders.

Supplementary Materials: The following are available online at http://www.mdpi.com/2076-2615/10/11/2002/s1, Table S1: Project and Sample ID for public access of the whole genome sequenced genomes used in this study. EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available. Figure S1: Neck skin lesions of the maternal-grandmother. Note the presence of bald areas and scabs.

Author Contributions: Conceptualization, M.B. and C.D.; methodology, J.G.P.J.; software, I.M.H. and A.L.; validation, C.B., I.M.B.V., and C.D.; formal analysis, I.M.H. and A.L.; investigation, J.G.P.J., I.M.B.V., A.L., and M.B.; resources, I.M.B.V., C.B., and C.D.; data curation, I.M.H.; writing—original draft preparation, J.G.P.J., I.M.H., and I.M.B.V.; writing—review and editing, J.G.P.J., I.M.H., I.M.B.V., A.L., C.B., M.B., and C.D.; visualization, J.G.P.J.

and I.M.B.V.; supervision, M.B. and C.D.; project administration, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by the Swiss National Science Foundation.

Acknowledgments: We acknowledge Stefano Allodi for referring the clinical cases. We thank Nathalie Besuchet Schmutz, Manuela Bozzo, Erika Bürgi, and Francesca Caravello for expert technical assistance and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure. Monika Welle is acknowledged for the case discussion.

Conflicts of Interest: C.B. is an editorial board member of Animals, but has not in any way been involved in or interacted with the journal's review process or editorial decision-making. The authors declare no conflict of interest.

References

1. Leeb, T.; Müller, E.J.; Roosje, P.; Welle, M. Genetic testing in veterinary dermatology. *Vet. Dermatol.* **2017**, *28*, e1. [CrossRef] [PubMed]
2. Malfait, F.; Francomano, C.; Byers, P.; Belmont, J.; Berglund, B.; Black, J.; Bloom, L.; Bowen, J.M.; Brady, A.F.; Burrows, N.P.; et al. The 2017 international classification of the Ehlers–Danlos syndromes. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **2017**, *175*, 8–26. [CrossRef] [PubMed]
3. Chiarelli, N.; Ritelli, M.; Zoppi, N.; Colombi, M. Cellular and molecular mechanisms in the pathogenesis of classical, vascular, and hypermobile ehlers-danlos syndromes. *Genes* **2019**, *10*, 609. [CrossRef] [PubMed]
4. Ritelli, M.; Dordoni, C.; Venturini, M.; Chiarelli, N.; Quinzani, S.; Traversa, M.; Zoppi, N.; Vascellaro, A.; Wischmeijer, A.; Manfredini, E.; et al. Clinical and molecular characterization of 40 patients with classic Ehlers-Danlos syndrome: Identification of 18 COL5A1 and 2 COL5A2 novel mutations. *Orphanet J. Rare Dis.* **2013**, *8*, 1–19. [CrossRef] [PubMed]
5. Monthoux, C.; de Brot, S.; Jackson, M.; Bleul, U.; Walter, J. Skin malformations in a neonatal foal tested homozygous positive for Warmblood Fragile Foal Syndrome. *BMC Vet. Res.* **2015**, *11*, 1–8. [CrossRef] [PubMed]
6. Tryon, R.C.; White, S.D.; Bannasch, D.L. Homozygosity mapping approach identifies a missense mutation in equine cyclophilin B (PPIB) associated with HERDA in the American Quarter Horse. *Genomics* **2007**, *90*, 93–102. [CrossRef]
7. Jaffey, J.A.; Bullock, G.; Teplin, E.; Guo, J.; Villani, N.A.; Mhlanga-Mutangadura, T.; Schnabel, R.D.; Cohn, L.A.; Johnson, G.S. A homozygous ADAMTS2 nonsense mutation in a Doberman Pinscher dog with Ehlers Danlos syndrome and extreme skin fragility. *Anim. Genet.* **2019**, *50*, 543–545. [CrossRef]
8. Bauer, A.; Bateman, J.F.; Lamandé, S.R.; Hanssen, E.; Kirejczyk, S.G.M.; Yee, M.; Ramiche, A.; Jagannathan, V.; Welle, M.; Leeb, T.; et al. Identification of two independent COL5A1 variants in dogs with ehlers-danlos syndrome. *Genes* **2019**, *10*, 731. [CrossRef]
9. Bauer, A.; de Lucia, M.; Leuthard, F.; Jagannathan, V.; Leeb, T. Compound heterozygosity for TNXB genetic variants in a mixed-breed dog with Ehlers-Danlos syndrome. *Anim. Genet.* **2019**, *50*, 546–549. [CrossRef]
10. Spycher, M.; Bauer, A.; Jagannathan, V.; Frizzi, M.; De Lucia, M.; Leeb, T. A frameshift variant in the COL5A1 gene in a cat with Ehlers-Danlos syndrome. *Anim. Genet.* **2018**, *49*, 641–644. [CrossRef]
11. Hegreberg, G.A.; Padgett, G.A.; Gorham, J.R.; Henson, J.B. A heritable connective tissue disease of dogs and Mink Resembling the Ehlers-Danlos. *J. Investig. Dermatol.* **1970**, *54*, 377–380. [CrossRef] [PubMed]
12. Harvey, R.; Brown, P.; Young, R.; Whitbread, T. A connective tissue defect in two rabbits similar to the Ehlers-Danlos syndrome. *Vet. Rec.* **1990**, *126*, 130–132. [PubMed]
13. Joller, S.; Berenguer Veiga, I.; Drögemüller, C. Dermatosparaxis in White Dorper sheep: Confirmation of a causative nonsense mutation in ADAMTS2. *Anim. Genet.* **2017**, *48*, 729–730. [CrossRef]
14. Zhou, H.; Hickford, J.G.H.; Fang, Q. A premature stop codon in the ADAMTS2 gene is likely to be responsible for dermatosparaxis in Dorper sheep. *Anim. Genet.* **2012**, *43*, 471–473. [CrossRef] [PubMed]
15. Monteagudo, L.V.; Ferrer, L.M.; Catalan-Insa, E.; Savva, D.; Mcguffin, L.J.; Tejedor, M.T. In silico identification and three-dimensional modelling of the missense mutation in ADAMTS2 in a sheep flock with dermatosparaxis. *Vet. Dermatol.* **2015**, *26*, 49–52. [CrossRef]
16. Carty, C.I.; Lee, A.M.; Wienandt, N.A.E.; Stevens, E.L.; Alves, D.A.; Browne, J.A.; Bryan, J.; Ryan, E.G.; Cassidy, J.P. Dermatosparaxis in two Limousin calves. *Ir. Vet. J.* **2016**, *69*, 1–5. [CrossRef]

17. Tajima, M.; Miyake, S.; Takehana, K.; Kobayashi, A.; Yamato, O.; Maede, Y. Gene defect of dermatan sulfate proteoglycan of cattle affected with a variant form of Ehlers-Danlos syndrome. *J. Vet. Intern. Med.* **1999**, *13*, 202–205. [CrossRef]
18. Colige, A.; Sieron, A.L.; Li, S.W.; Schwarze, U.; Petty, E.; Wertelecki, W.; Wilcox, W.; Krakow, D.; Cohn, D.H.; Reardon, W.; et al. Human ehlers-danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *Am. J. Hum. Genet.* **1999**, *65*, 308–317. [CrossRef]
19. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, 1–9. [CrossRef]
20. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef]
21. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef] [PubMed]
22. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
23. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant review with the integrative genomics viewer. *Cancer Res.* **2017**, *77*, e31–e34. [CrossRef]
24. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef]
25. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.; Lin, G.N.; Nam, H.-J.; Mort, M.; Cooper, D.; Sebat, J.; Iakoucheva, L.; et al. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv* **2017**, 134981. [CrossRef]
26. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Comput. Biol.* **2014**, *10*, 1–11. [CrossRef] [PubMed]
27. Dalgleish, R. The Human Collagen Mutation Database 1998. *Nucleic Acids Res.* **1998**, *26*, 253–255. [CrossRef] [PubMed]
28. Dalgleish, R.; Teh, W.K. Ehlers Danlos Syndrome Variant Database. Available online: https://eds.gene.le.ac.uk/home.php?select_db=COL5A1 (accessed on 29 September 2020).
29. Symoens, S.; Syx, D.; Malfait, F.; Callewaert, B.; de Backer, J.; Vanakker, O.; Coucke, P.; de Paepe, A. Comprehensive molecular analysis demonstrates type V collagen mutations in over 90% of patients with classic EDS and allows to refine diagnostic criteria. *Hum. Mutat.* **2012**, *33*, 1485–1493. [CrossRef]
30. Zoppi, N.; Chiarelli, N.; Cinquina, V.; Ritelli, M.; Colombi, M. GLUT10 deficiency leads to oxidative stress and non-canonical αvβ3 integrin-mediated TGFβ signalling associated with extracellular matrix disarray in arterial tortuosity syndrome skin fibroblasts. *Hum. Mol. Genet.* **2015**, *24*, 6769–6787. [CrossRef]
31. Richards, A.J.; Martin, S.; Nicholls, A.C.; Harrison, J.B.; Pope, F.M.; Burrows, N.P. A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type II. *J. Med. Genet.* **1998**, *35*, 846–848. [CrossRef]
32. Zaffino Heyerhoff, J.C.; LeBlanc, S.J.; DeVries, T.J.; Nash, C.G.R.; Gibbons, J.; Orsel, K.; Barkema, H.W.; Solano, L.; Rushen, J.; de Passillé, A.M.; et al. Prevalence of and factors associated with hock, knee, and neck injuries on dairy cows in freestall housing in Canada. *J. Dairy Sci.* **2014**, *97*, 173–184. [CrossRef] [PubMed]
33. Mitchell, A.L.; Schwarze, U.; Jennings, J.F.; Byers, P.H. Molecular Mechanisms of Classical Ehlers-Danlos Syndrome. *Hum. Mutat.* **2009**, *30*, 995–1002. [CrossRef]
34. Lamandé, S.R.; Bateman, J.F. Genetic Disorders of the Extracellular Matrix. *Anat. Rec.* **2020**, *303*, 1527–1542. [CrossRef] [PubMed]
35. DE, B. Type V collagen: Heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron* **2001**, *32*, 223–237.

A *de novo* mutation in *KRT5* in a crossbred calf with epidermolysis bullosa simplex

Journal:	Journal of Veterinary Internal Medicine
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.1111/jvim.15943

Received: 3 July 2020 | Accepted: 13 October 2020

DOI: 10.1111/jvim.15943

CASE REPORT

Journal of Veterinary Internal Medicine Open Access

ACVIM American College of Veterinary Internal Medicine

A de novo mutation in *KRT5* in a crossbred calf with epidermolysis bullosa simplex

Joana G. P. Jacinto[1,2] | Irene M. Häfliger[2] | Inês M. B. Veiga[3] | Cord Drögemüller[2] | Jørgen S. Agerholm[4]

[1]Department of Veterinary Medical Sciences, University of Bologna, Bologna, Italy

[2]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[3]Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[4]Department of Veterinary Clinical Sciences, University of Copenhagen, Copenhagen, Denmark

Correspondence
Cord Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern. Bremgartenstr. 109a, 3001 Bern, Switzerland.
Email: cord.droegemueller@vetsuisse.unibe.ch

Abstract

A 6-day-old Belgian Blue-Holstein calf was referred because of a syndrome resembling epidermolysis bullosa simplex (EBS). The clinical phenotype included irregular and differently sized erosions and ulcerations spread over the body, in particular on the limbs and over bone prominences, as well as in the nasal planum and oral mucosa. Blisters were easily induced by rubbing the skin. The skin lesions displayed a clear dermal-epidermal separation at the level of the basal cell layer. Post mortem examination revealed erosions in the pharynx, proximal esophagus, and rumen. Whole-genome sequencing revealed a heterozygous disruptive in-frame deletion variant in *KRT5* (c.534_536delCAA). Genotyping of both parents confirmed the variant as de novo mutation. Clinicopathological and genetic findings were consistent with the diagnosis of *KRT5*-related EBS providing the second example of a spontaneous mutation causing epidermolysis bullosa in cattle.

KEYWORDS

cattle, KRT5, precision medicine, skin fragility, WGS

1 | INTRODUCTION

Epidermolysis bullosa (EB) encompasses a heterogeneous group of genetic mechanobullous disorders characterized by blistering from even minor mechanical trauma with disruption at the dermal-epidermal junction.[1] Epidermolysis bullosa disorders are characterized by clinical heterogeneity, both in their appearance and severity. The disease might be congenital or develop later in life. In congenital cases, the lesions are more severe, accompanied by mucosal fragility, and might involve other organs than the skin.[2] In noncongenital cases, the skin fragility is less severe and the lesions are usually localized to the extremities of the limbs, occasionally only expressed as nail/hoof dystrophy.[3] Based on the ultrastructural level of skin cleavage, there are 4 major classical types: EB simplex (EBS), junctional EB (JEB), dystrophic EB (DEB), and Kindler EB (KEB).[4] In EBS, the skin cleavage occurs within the epidermis, in JEB within the lamina lucida and in DEB within the superficial dermis. Kindler EB might present with all 3 cleavage levels.[3] The same classification might be used in veterinary medicine; however, cases of KEB are not reported in domestic animals. In human, these phenotypical classifications are complicated by the fact that, depending on the variant, the same gene might be associated with different modes of inheritance, thus resulting in distinct clinical phenotypes.[5] Dystrophic EB and EBS phenotypes might be inherited either dominantly or recessively, and might be caused by pathogenic

Abbreviations: DEB, dystrophic EB; EBS, EB simplex; EDTA, ethylenediaminetetraacetic acid; EB, epidermolysis bullosa; HE, hematoxylin and eosin; HIM, helix initiation peptide motif; HTM, helix termination peptide motif; IGV, Integrative Genomics Viewer; JEB, junctional EB; KRT5, keratin 5; KEB, Kindler EB; WGS, whole-genome sequencing.

variants in different genes.[4] In human medicine, EB is associated with more than 1000 variants in at least 18 genes encoding structural proteins.[1,5]

Four EB-related causative recessive variants are known in cattle,[6-9] 3 in dogs,[10-13] 2 in sheep,[14,15] and 2 in horses,[16-19] and 1 dominant variant is known in cattle[20] (Table S1). For cattle, a dominant form of EBS is associated with a keratin 5 (*KRT5*) missense variant (OMIA 002081-9913),[20] and recessive forms of JEB are associated with deleterious variants in *ITGB4* (OMIA 001948-9913), *LAMA3* (OMIA 001677-9913), and *LAMC2* (OMIA 001678-9913).[6-8] In addition, a form of recessive DEB is associated with a nonsense variant in *COL7A1* (OMIA 000341-991).[9]

Older reports of familial occurrence of EB,[21] outbreaks of several inherited-related cases in single herds[21-23] and sporadic cases of EB[24-27] in cattle exist. In the abovementioned cases,[21-23] the diagnosis was based only on the clinical and histopathological findings.

The majority of the previous reports focused on disorders with a recessive inheritance. However, single cases because of dominant acting de novo variants might occur sporadically without impact on breeding. At present, this obvious genetic heterogeneity could be analyzed in cattle using whole-genome sequencing (WGS)-based precision diagnostics.[28] Therefore, the purpose of this study was to characterize the clinical and pathological phenotype of an EBS-affected calf, and to evaluate its possible genetic etiology using WGS.

2 | CASE DESCRIPTION

A 6-day-old (46 kg) male Belgian Blue-Holstein crossbred calf was submitted for clinical investigation because of ulcerations of the skin and nasal planum since short time after birth. The animal was delivered after a gestation period of 287 days.

The cutaneous lesions were characterized by widespread irregular erosions and ulcerations of various sizes on most parts of the body (Figure 1A), but in particular on the limbs (Figure 1B) and over bony prominences. Upon handling, the epidermis easily separated leaving a blister with a black colored, nonhemorrhagic base indicating a separation superficial to the stratum basale. Peracute blister were easily induced by rubbing the skin by an eraser after having cut the hair locally. Older lesions consisted of ulcerations covered by crusts and occasional acute hemorrhage. On the nasal platum, lips and nares extensive ulcerations were present; the calf also showed a purulent nasal discharge (Figure 1C). Moreover, the animal seemed to be in pain when walking on a hard surface. The aspect resembled EB and therefore was further referred to the Danish surveillance program for genetic diseases in cattle for further examination. Because of the poor prognosis and the painful situation, the calf was euthanized for welfare reasons by IV administration of an overdose of pentobarbital. In addition to the skin lesions, gross pathologic examination revealed erosions in the oral cavity, pharynx, proximal esophagus, and rumen.

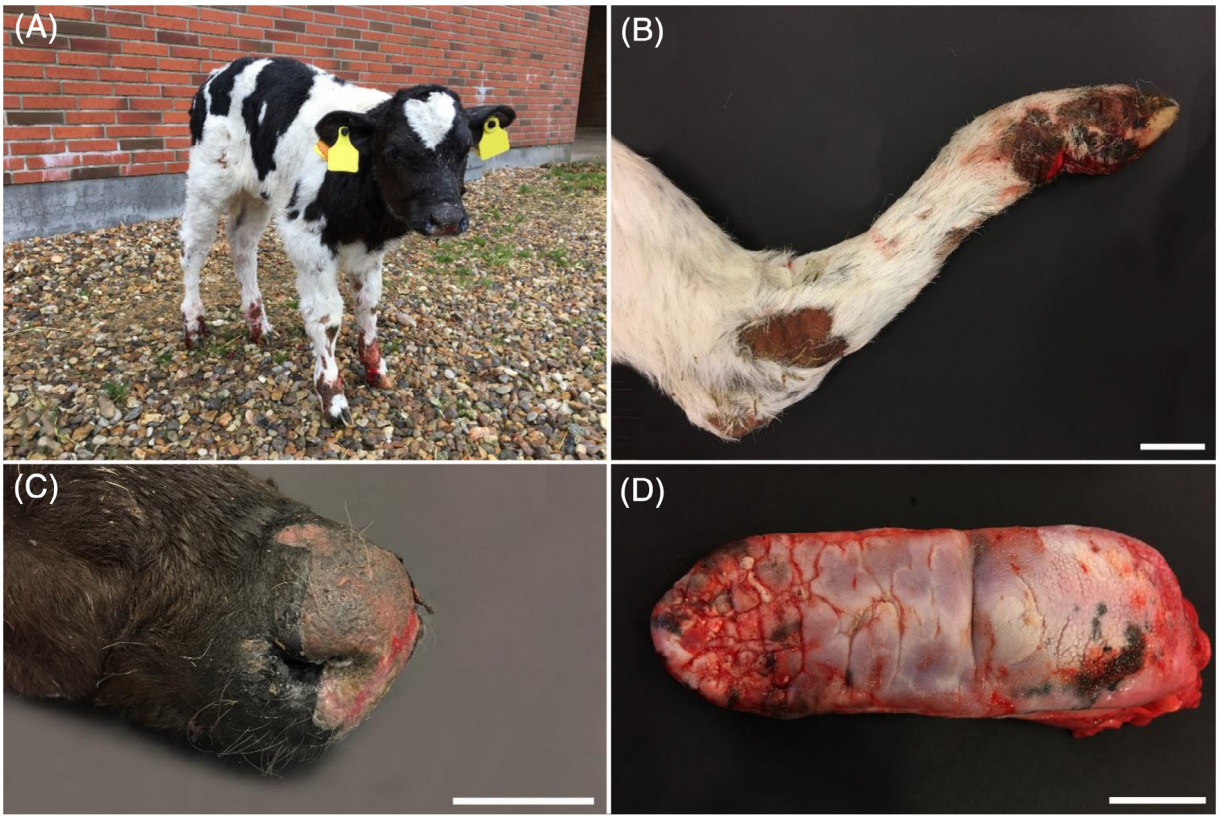


FIGURE 1 Lesions in the affected calf. A, Widespread irregular ulcerations of various sizes on most parts of the body. B, Irregular ulceration on the hindlimb. Scale bar = 5 cm. C, Extensive ulcerations on the nasal planum, lips and nares; also, purulent nasal discharge was present. Scale bar = 2.5 cm. D, Thickened and the furrows of the epithelium on the dorsal surface of the tongue. Scale bar = 5 cm

The epithelium on the dorsal surface of the tongue was thickened and with furrows (Figure 1D). The incisor teeth were disorganized and not completely erupted and the surrounding parts of the mandibles appeared thickened and cystic. The hoofs seemed intact, yet when sawed longitudinally, the capsule was partly separated from the dermal lamella with hemorrhage in the interface.

Immediately before euthanasia, the skin covering the dorsal part of the pelvis was gently scrubbed with an eraser with blister formation. Skin biopsies from this area and from other representative cutaneous lesions were taken immediately after euthanasia for histological analysis, whereas the necropsy was completed at the university a few hours later. Additional specimens for histological analysis were then collected, including the oral mucosa, pharynx, rumen, reticulum, and major internal organs. All collected samples were fixed in 10% neutral buffered formalin, trimmed, processed, embedded in paraffin wax, sectioned at 4 to 5 μm, and stained by hematoxylin and eosin (HE). Histologically, the peracute lesions induced by rubbing before euthanasia displayed a very striking, multifocal to coalescing dermal-epidermal separation at the level of the basal layer, which extended into the wall of the hair follicle infundibula (Figure 2). The spontaneously occurring, chronic lesions present in the nasal planum and in the distal limbs displayed a multifocal to coalescing epithelial loss with consequent severe ulceration and underlying neutrophilic infiltration, replacement of the papillary dermis by granulation tissue, and re-epithelialization. A multifocal dermal-epidermal separation at the level of the basal cell layer with multifocal underlying accumulation of free erythrocytes and fibrin exudation was occasionally visible at the border of the ulcerated areas.

In the tongue, the mucosa of the dorsal surface showed a marked parakeratotic hyperkeratosis. At the lateral borders, where the epithelium had a normal thickness, areas with complete loss of mucosa were observed. The superficial layer of the submucosal connective tissue had a necrotic surface, intense hyperemia, and infiltration with neutrophils. Mucosa cleavage in the adjacent areas was not observed, but the height of the epithelium gradually decreased. The stratum spinosum showed ballooning degeneration and in these areas, the stratum corneum was not present. In the pharyngeal lining, multiple intensely inflamed ulcers covered by a debris of fibrin, degenerated neutrophils, erythrocytes, bacterial colonies were present.

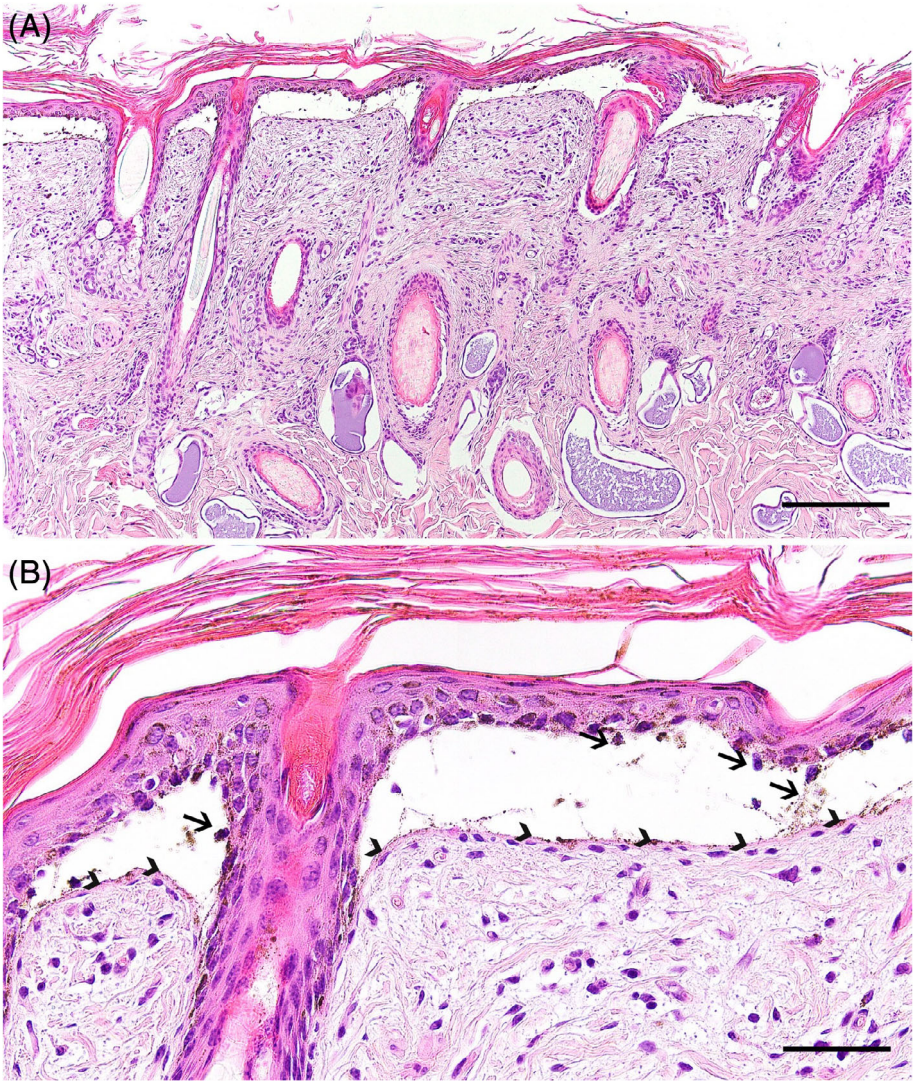


FIGURE 2 Histopathological findings of the affected calf. A, The peracute cutaneous lesions displayed an extensive dermal-epidermal separation at the level of the basal layer, which also affected the wall of the hair follicle infundibula. HE staining, scale bar = 200 μm. B, Higher magnification of the dermal-epidermal detachment at the basal cell layer level, where it is possible to observe the presence of occasional basal cell remains within the cleft (arrows) and the basement membrane overlying the dermis (arrowheads). HE staining, scale bar = 50 μm

The epithelium of the rumen and reticulum was normally developed but an acute suppurative multifocal rumenitis was present. Separation of the epithelium from the underlying submucosa was seen in some areas but considered as a post mortem artifact. Additional findings included suppurative periportal hepatitis and multifocal intense pulmonary hyperemia associated with fibrin in the alveoli. Other tissues were unremarkable. The histopathological findings in the skin and pharyngeal linings resembled EBS.

Additionally, WGS using the Illumina NovaSeq6000 was performed on DNA extracted from ethylenediaminetetraacetic acid (EDTA) blood of the calf. The sequenced reads were mapped to the ARS-UCD1.2 reference genome resulting in an average read depth of approximately 19×,[29] and single-nucleotide variants and small indel variants were called. The applied software and steps to process fastq files into binary alignment map and genomic variant call format files were in accordance with the 1000 Bull Genomes Project processing guidelines of run 7 (www.1000bullgenomes.com),[30] except for the trimming, which was performed using fastp.[31] Further preparation of the genomic data had been done according to Häfliger et al.[32] In order to find private variants, we compared the genotypes of the affected calf with 493 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that are publicly available (Table S2) in the European Nucleotide Archive (SAMEA6528898 is the sample accession number of the case; http://www.ebi.ac.uk/en). Integrative Genomics Viewer (IGV)[33] software was used for visual inspection of candidate variants. A total of 115 private protein-changing variants with a moderate or high predicted impact on the encoded protein, located within 108 different genes or loci, were identified. These variants were further checked for their occurrence in a global control cohort of 3103 genomes of a variety of breeds (1000 Bull Genomes Project run 7; www.1000bullgenomes.com), which revealed 26 protein-changing variants exclusively present heterozygous in the genome of the EBS-affected calf. These 26 variants located within 25 different genes or loci (Table S3) were subsequently visually inspected using IGV software confirming all as true variants. Of all these 26 remaining private variants, only 1 occurred in a candidate for EB: keratin 5 (*KRT5*). The variant was a heterozygous disruptive in-frame deletion variant on chromosome 5: 27367604delCAA (NM_001008663.1:c.534_536delCAA), leading to a loss of an asparagine amino acid at residue 178 of the encoded *KRT5* protein (NP_001008663.1:p.Asn178del). This variant affecting an EB candidate gene was further investigated as likely causal mutation for the observed phenotype.

To confirm that the c.534_536delCAA variant in *KRT5* was a de novo mutation, the affected genomic region was amplified by polymerase chain reaction (PCR) and Sanger sequenced in the affected calf, its Belgian Blue sire and Holstein dam based on DNA extracted from EDTA blood of the dam, and from both EDTA blood and semen of the sire. PCR products were amplified using flanking primers for the *KRT5* exon 1 deletion with 5′-AGGCATCCAAGAGGTCACCG-3′ (forward primer) and 5′-TAGCACATATCCCACACTCATGG-3′ (reverse primer). Sequence data were analyzed using Sequencher 5.1 (GeneCodes). Analyzing the sequencing data, we concluded that only the EBS-affected calf was heterozygous for the *KRT5* variant and the dam and sire were both homozygous for the wild type allele in all analyzed samples including both semen and blood of the sire (Figure 3).

3 | DISCUSSION

The clinical and pathological findings in the calf were consistent with EB. Although most lesions had the appearance of unspecific inflamed ulcerations, which usually develop shortly after dermal-epidermal separation, blistering could be easily induced in intact skin by rubbing the skin surface. Histopathological analysis of these lesions showed that the dermal-epidermal separation occurred at the level of the basal cell layer, which is suggestive for EBS.

The predicted deleterious protein effect c.534_536delCAA variant and the conservation of the affected asparagine amino acid residue at position 178 in the helix initiation motif (HIM) of the highly conserved 1A rod domain of *KRT5* suggest that this de novo mutation variant is certainly pathogenic (Figure 4). The mutation most likely occurred post-zygotically during the calf's fetal development as it was absent in both parents. In cattle, a de novo missense variant in *KRT5* is reported in an asymptomatic Friesian-Jersey crossbred mosaic sire and EBS-affected offspring. That mutation results in an amino acid exchange (p.Glu478Lys) in the final glutamic acid of the KLLEGE motif of the highly conserved 2B rod domain of *KRT5* (Figure 4).[20]

The early onset combined with multifocal to widespread lesions in the skin and mucosal membranes classifies this condition as a severe form of EB. The reported bovine *KRT5* associated EB case also had such lesions[20]; this indicates that mutations in the *KRT5* in cattle might cause severe EBS when present. Histologically the present case showed suppurative rumenitis and rumen epithelial detachment. Because of the delayed necropsy, the latter might have been a post mortem artifact. However, no ulcerations were observed in the forestomach and inflamed areas were covered by an epithelium, therefore a possible association with EB remains hypothetical.

The case of EB presented in this study can be classified as EBS. In humans, among the several subtypes reported in the literature, the most common EBS subtypes might be considered the so-called localized, severe, and intermediate form. In the localized EBS, the blisters are present only on the extremities of the limbs. Rare phenotypical subtypes of localized EBS are associated with nephropathy.[34] In severe forms of EBS the lesions are present from birth, are more diffuse on the body, with main severity on the extremities and over bone prominences, and deep ulceration might be observed. After some time, large tense blisters might arise spontaneously or secondary to minor trauma. The blisters characteristically have an arciform pattern and eventuate with crusts to necrosis with a visual similarity to inflammatory plaques.[1] Also, the oral mucosa might be affected. Some phenotypical subtypes of the severe EBS are accompanied by other systemic complications with subsequent growth retardation, nutritional deficiency or even lethal outcome because of secondary infections or respiratory failure.[35] In intermediate EBS the skin lesions are diffuse on the body, however they are not so serious as in the severe

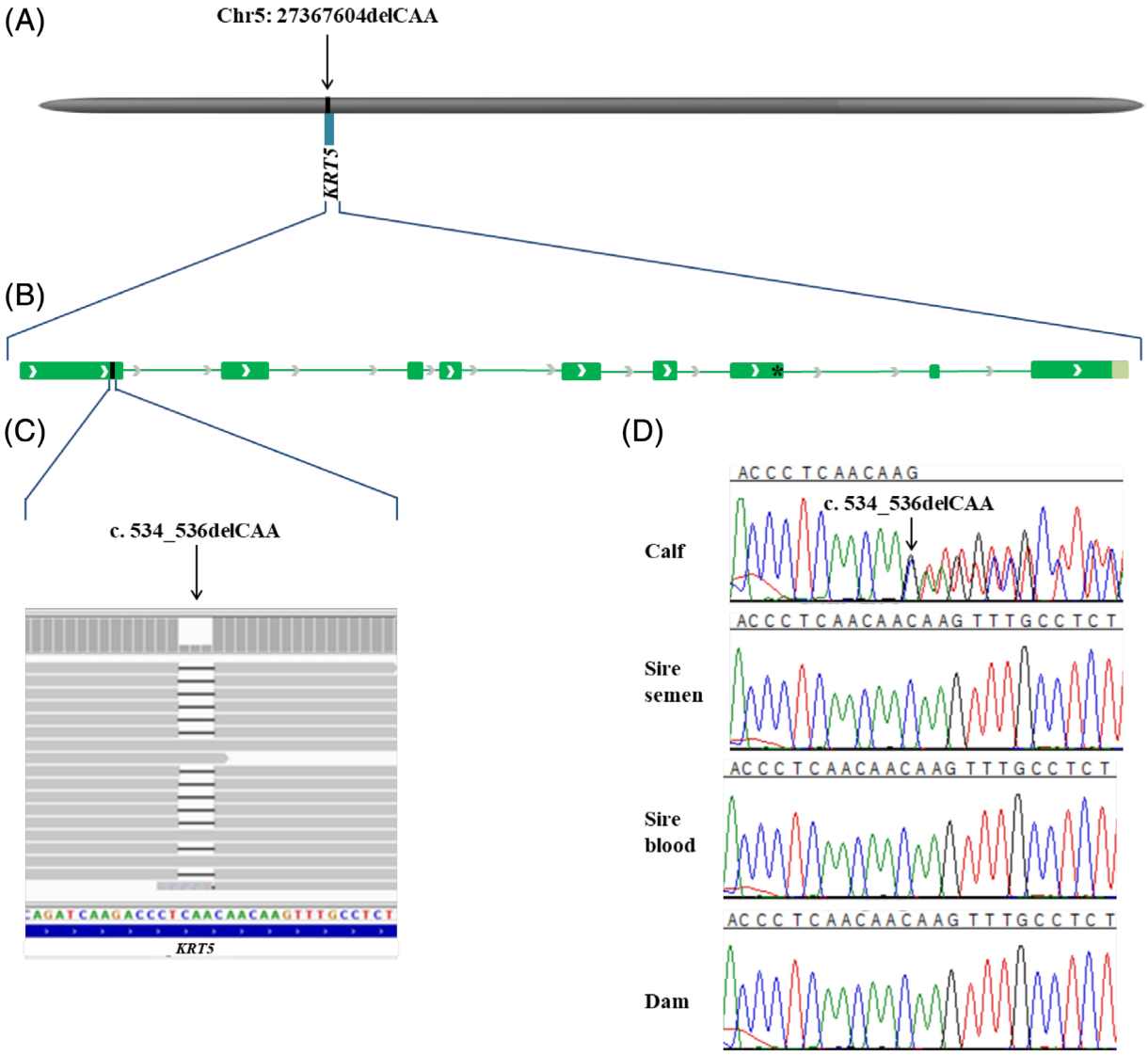


FIGURE 3 Schematic diagram of the *KRT5* gene showing the location of the candidate causal variant NM_001008663.1:c.534_536delCAA of the affected calf. A, Location of the bovine *KRT5* gene, Chr5:27 367 078-27 372 929 and causal variant, Chr5:27 367 604 on the ARS-UCD1.2 bovine genome assembly. B, Genomic structure of *KRT5* gene. Green boxes represent the exons. The c.534_536delCAA is located in the 1st exon of KRT5 gene. The previous reported g.27371128G>A missense variant is located in the 7th exon of *KRT5* gene. C, Integrative Genomics Viewer (IGV) screenshot presenting the *KRT5* variant. D, Sanger sequencing results confirmed that the variant occurred de novo as sequencing of PCR products from DNA of both parents (for the sire both semen and blood) showed that the variant was absent

EBS. Rare phenotypical subtypes of intermediate EBS are accompanied by cardiomyopathy[36,37] and muscular dystrophy.[38,39]

In human, intermediate EBS with muscular dystrophy is often associated with enamel hypoplasia.[39] Furthermore, beyond the characteristic skin lesions, it might include diffuse alopecia, short stature, slow weight gain, punctate keratitis, urethral strictures, muscular dystrophy, and degenerative changes with increased connective tissue.[35] Such findings were not observed in the studied calf.

In human medicine, a de novo missense variant in *KRT5* resulting in an amino acid exchange (p.Asn177Ser) in the HIM has been reported in a patient with localized EBS.[40] This human KRT5 protein position corresponds to the position of the p.Asn178del variant present in this case. In human medicine, the localized, severe, and intermediate subtypes of the EBS are mostly linked to an autosomal dominant pattern and are associated to a high rate of de novo mutations. The mutation in the present case was in that aspect similar to the human cases, in that it was autosomal dominant and of de novo origin. The most common mutations are caused by monoallelic pathogenic missense, nonsense, frameshift, or splice site variants or in frame deletions in *KRT5* and *KRT14*.[41] However, rare EBS subtypes might be associated with pathogenic variants in other genes, such as *EXPH5*,[42] *KLHL24*,[43] *DST*,[44] *PLEC*,[45-48] and *CD151*.[34]

Epidermolysis bullosa simplex is a rare disorder known in man and animals. Rare disorders such as EBS in livestock are usually not

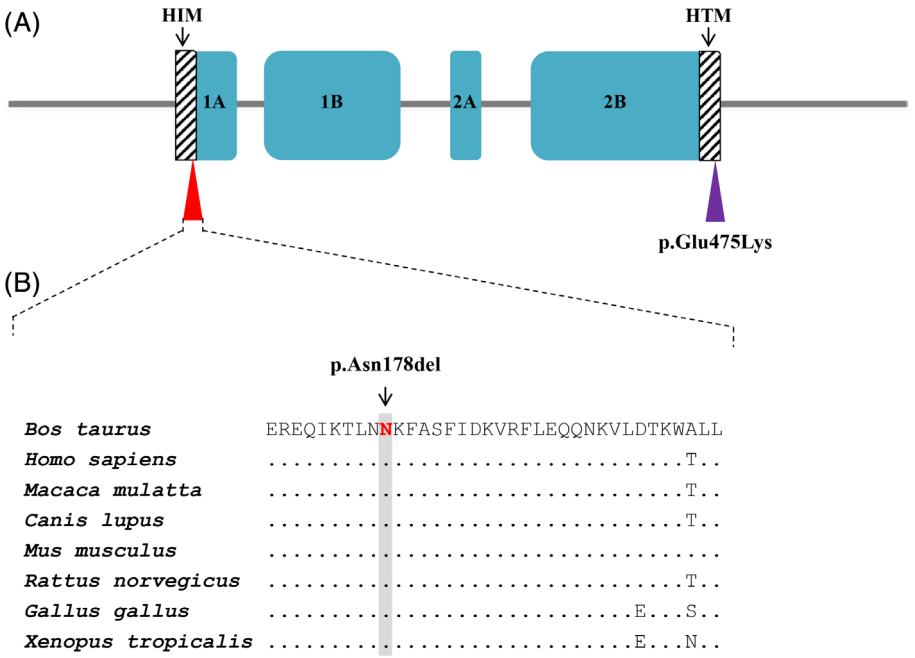


FIGURE 4 Schematic diagram of the KRT5 protein. A, Domain and region information of KRT5 protein obtained from the UniProt database (http://www.uniprot.org/; accession number: Q5XQN5). The p.Asn178del variant in the helix initiation peptide motif (HIM) is indicated by the red triangle. The EBS-causing p.Glu475Lys variant previously reported[20] in the helix termination peptide motif (HTM) is indicated by a violet triangle. The arrows indicate the HIM and the HTM. B, Multiple sequence alignment of 1A rod domain of the of KRT5 protein encompassing the region of the p.Asn178del variant demonstrates a complete evolutionary conservation across species. The observed variant is indicated by an arrow and the respective position highlighted in gray. Protein sequences accession numbers in NCBI for each species are NP_001008663.1 (*Bos taurus*), NP_000415.2 (*Homo sapiens*), XP_002798641.1 (*Macaca mulatta*), XP_005636850.1 (*Canis familiaris*), NP_081287.1 (*Mus musculus*), NP_899162.1 (*Rattus norvegicus*), NP_001001195.1 (*Gallus gallus*), and NP_001072377.1 (*Xenopus tropicalis*)

diagnosed to the molecular level, mainly because of lack of resources and diagnostic tools as well as low value and often-short lifespan of the animals. The report of this case allowed the performance of a complete clinical, pathological, and molecular genetic study enabling the diagnosis of a severe form of EBS. Furthermore, this example highlights the utility of WGS-based precision diagnostics for understanding rare disorders in animals with an available reference genome sequence and the value of surveillance of cattle breeding populations for harmful genetic disorders.

ACKNOWLEDGMENTS

No funding was received for this study. The authors acknowledge the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome sequencing and the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure. The whole-genome sequencing data are available under the study accession no. PRJEB18113 at the European Nucleotide Archive (www.ebi.ac.uk/ena; sample accessions SAMEA6528898). Monika Welle from the Institute of Veterinary Pathology of the University of Bern is acknowledged for the fruitful case discussion.

CONFLICT OF INTEREST DECLARATION

Authors declare no conflicts of interests.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label use of abtimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC)

This study was not based on an invasive animal experiment but was based on a spontaneously occurring case; therefore, there are no associated permit numbers.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Joana G. P. Jacinto https://orcid.org/0000-0002-6438-7975
Irene M. Häfliger https://orcid.org/0000-0002-5648-963X
Inês M. B. Veiga https://orcid.org/0000-0001-8554-790X
Cord Drögemüller https://orcid.org/0000-0001-9773-522X
Jørgen S. Agerholm https://orcid.org/0000-0003-1653-4552

REFERENCES

1. Fine JD, Bruckner-Tuderman L, Eady RAJ, et al. Inherited epidermolysis bullosa: updated recommendations on diagnosis and classification. *J Am Acad Dermatol*. 2014;70(6):1103-1126. https://doi.org/10.1016/j.jaad.2014.01.903.

2. Has C, Liu L, Bolling MC, et al. Clinical practice guidelines for laboratory diagnosis of epidermolysis bullosa. *Br J Dermatol*. 2020;182(3):574-592.
3. Has C, Fischer J. Inherited epidermolysis bullosa: new diagnostics and new clinical phenotypes. *Exp Dermatol*. 2019;28(10):1146-1152.
4. Has C, Bauer JW, Bodemer C, et al. Consensus reclassification of inherited epidermolysis bullosa and other disorders with skin fragility. *Br J Dermatol*. 2020;183(4):1-14. https://doi.org/10.1111/bjd.18921.
5. Has C, Bruckner-Tuderman L. The genetics of skin fragility. *Annu Rev Genomics Hum Genet*. 2014;15:245-268.
6. Peters M, Reber I, Jagannathan V, Raddatz B, Wohlsein P, Drögemüller C. DNA-based diagnosis of rare diseases in veterinary medicine: a 4.4 kb deletion of ITGB4 is associated with epidermolysis bullosa in Charolais cattle. *BMC Vet Res*. 2015;11(1):1-8.
7. Sartelet A, Harland C, Tamma N, et al. A stop-gain in the laminin, alpha 3 gene causes recessive junctional epidermolysis bullosa in Belgian Blue cattle. *Anim Genet*. 2015;46(5):566-570.
8. Murgiano L, Wiedemar N, Jagannathan V, Isling LK, Drögemüller C, Agerholm JS. Epidermolysa bullosa in Danish Hereford calves is caused by a deletion in LAMC2 gene. *BMC Vet Res*. 2015;11(1):23.
9. Menoud A, Welle M, Tetens J, Lichtner P, Drögemüller C. A COL7A1 Mutation Causes Dystrophic Epidermolysis Bullosa in Rotes Höhenvieh Cattle. *PLoS ONE*. 2012;7(6):e38823. http://dx.doi.org/10.1371/journal.pone.0038823.
10. Mauldin EA, Wang P, Olivry T, Henthorn PS, Casal ML. Epidermolysis bullosa simplex in sibling Eurasier dogs is caused by a PLEC non-sense variant. *Vet Dermatol*. 2017;28(1):10-e3.
11. Capt A, Spirito F, Guaguere E, Spadafora A, Ortonne JP, Meneguzzi G. Inherited junctional epidermolysis bullosa in the German pointer: establishment of a large animal model. *J Invest Dermatol*. 2005;124(3):530-535. https://doi.org/10.1111/j.0022-202X.2004.23584.x.
12. Niskanen J, Dillard K, Arumilli M, et al. Nonsense variant in COL7A1 causes recessive dystrophic epidermolysis bullosa in Central Asian Shepherd dogs. *PLoS One*. 2017;12(5):1-10.
13. Baldeschi C, Gache Y, Rattenholl A, et al. Genetic correction of canine dystrophic epidermolysis bullosa mediated by retroviral vectors. *Hum Mol Genet*. 2003;12(15):1897-1905.
14. Suárez-Vega A, Gutiérrez-Gil B, Benavides J, et al. Combining GWAS and RNA-Seq approaches for detection of the causal mutation for hereditary junctional epidermolysis bullosa in sheep. *PLoS One*. 2015;10(5):e0126416.
15. Mömke S, Kerkmann A, Wöhlke A, et al. A frameshift mutation within LAMC2 is responsible for herlitz type junctional epidermolysis bullosa (HJEB) in black headed mutton sheep. *PLoS One*. 2011;6(5):e18943.
16. Graves KT, Henney PJ, Ennis RB. Partial deletion of the LAMA3 gene is responsible for hereditary junctional epidermolysis bullosa in the American Saddlebred Horse. *Anim Genet*. 2009;40(1):35-41.
17. Cappelli K, Brachelente C, Passamonti F, Flati A, Silvestrelli M, Capomaccio S. First report of junctional epidermolysis bullosa (JEB) in the Italian draft horse. *BMC Vet Res*. 2015;11(1):2-5.
18. Milenkovic D, Chaffaux S, Taourit S, Guérin G. A mutation in the LAMC2 gene causes the Herlitz junctional epidermolysis bullosa (H-JEB) in two French draft horse breeds. *Genet Sel Evol*. 2003;35:249-256.
19. Spirito F, Charlesworth A, Linder K, Ortonne JP, Baird J, Meneguzzi G. Animal models for skin blistering conditions: absence of laminin 5 causes hereditary junctional mechanobullous disease in the Belgian horse. *J Invest Dermatol*. 2002;119(3):684-691. https://doi.org/10.1046/j.1523-1747.2002.01852.x.
20. Ford CA, Stanfield AM, Spelman RJ, et al. A mutation in bovine keratin 5 causing epidermolysis bullosa simplex, transmitted by a mosaic sire. *J Invest Dermatol*. 2005;124(6):1170-1176. https://doi.org/10.1111/j.0022-202X.2005.23610.x.
21. Bassett H. A congenital bovine epidermolysis resembling epidermolysis bullosa simplex of man. *Vet Rec*. 1987;121:8-11.
22. Thompson KG, Crandell RA, Rugeley WW, Sutherland RJ. A mechanobullous disease with sub-basilar separation in Brangus calves. *Vet Pathol*. 1985;22:283-285.
23. Foster AP, Skuse AM, Higgins RJ, et al. Epidermolysis Bullosa in calves in the United Kingdom. *J Comp Pathol*. 2010;142(4):336-340. https://doi.org/10.1016/j.jcpa.2009.10.001.
24. Deprez P, Maenhout T, De Cock H, et al. Epidermolysis bullosa in a calf: a case report [in Dutch]. *Vlaams Diergeneeskd Tijdschr*. 1993;62:155-159.
25. Agerholm JS. Congenital generalized epidermolysis bullosa in a calf. *Zentralbl Veterinarmed A*. 1994;41:139-142.
26. Stocker H, Lott G, Straumann U, Rüsch P. Epidermolysis bullosa in a calf [in German]. *Tierarztl Prax*. 1995;23:123-126.
27. Medeiros GX, Riet-Correa F, Armién AG, Dantas AFM, de Galiza GJN, Simões SVD. Junctional epidermolysis bullosa in a calf. *J Vet Diagnostic Investig*. 2012;24(1):231-234.
28. Bourneuf E, Otz P, Pausch H, et al. Rapid discovery of de novo deleterious mutations in cattle enhances the value of livestock as model species. *Sci Rep*. 2017;7(1):1-19.
29. Rosen BD, Bickhart DM, Schnabel RD, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020;9(3):1-9.
30. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. *Annu Rev Anim Biosci*. 2019;7:89-102.
31. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890.
32. Häfliger IM, Wiedemar N, Švara T, et al. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim Genet*. 2020;51(3):382-390.
33. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the integrative genomics viewer. *Cancer Res*. 2017;77(21):e31-e34.
34. Crew VK, Burton N, Kagan A, et al. CD151, the first member of the tetraspanin (TM4) superfamily detected on erythrocytes, is essential for the correct assembly of human basement membranes in kidney and skin. *Blood*. 2004;104(8):2217-2223.
35. Fine JD, Johnson LB, Weiner M, Suchindran C. Cause-specific risks of childhood death in inherited epidermolysis bullosa. *J Pediatr*. 2008;152(2):276-280.e2.
36. Schwieger-Briel A, Fuentes I, Castiglia D, et al. Epidermolysis bullosa simplex with KLHL24 mutations is associated with dilated cardiomyopathy. *J Invest Dermatol*. 2019;139(1):244-249.
37. Bolling MC, Pas HH, De Visser M, et al. PLEC1 mutations underlie adult-onset dilated cardiomyopathy in epidermolysis bullosa simplex with muscular dystrophy. *J Invest Dermatol*. 2010;130(4):1178-1181.
38. Kyrova J, Kopeckova L, Buckova H, et al. Epidermolysis bullosa simplex with muscular dystrophy. Review of the literature and a case report. *J Dermatol Case Rep*. 2016;10(3):39-48.
39. OMIM. Online Mendelian Inheritance in Man: phenotype 226670 Epidermolysis bullosa simplex with muscular dystrophy; EBSMD [Internet]. https://omim.org/entry/226670?search=epidermolysis bullosa simplex with teeth alterations&highlight=alteration bullosa epidermolysi simplex tooth with. Accessed March 30, 2020.
40. Liovic M, Bowden PE, Marks R, Komel R. A mutation (N177S) in the structurally conserved helix initiation peptide motif of keratin 5 causes a mild EBS phenotype. *Exp Dermatol*. 2004;13(5):332-334.
41. Szeverenyi I, Cassidy AJ, Cheuk WC, et al. The human intermediate filament database: comprehensive information on a gene family involved in many human diseases. *Hum Mutat*. 2008;29(3):351-360.
42. Turcan I, Pasmooij AMG, Van Den Akker PC, Lemmink H, Sinke RJ, Jonkman MF. Association of epidermolysis bullosa simplex with mottled pigmentation and EXPH5 mutations. *JAMA Dermatol*. 2016;152(10):1137-1141.

43. Lee JYW, Liu L, Hsu CK, et al. Mutations in KLHL24 add to the molecular heterogeneity of epidermolysis bullosa simplex. *J Invest Dermatol*. 2017;137(6):1378-1380.
44. Groves RW, Liu L, Dopping-Hepenstal PJ, et al. A homozygous nonsense mutation within the dystonin gene coding for the coiled-coil domain of the epithelial isoform of BPAG1 underlies a new subtype of autosomal recessive epidermolysis bullosa simplex. *J Invest Dermatol*. 2010;130(6):1551-1557. https://doi.org/10.1038/jid.2010.19.
45. Koss-Harnes D, Høyheim B, Anton-Lamprecht I, et al. A site-specific plectin mutation causes dominant epidermolysis bullosa simplex Ogna: two identical de novo mutations. *J Invest Dermatol*. 2002;118(1):87-93.
46. Smith FJ, Eady RA, Leigh IM, et al. Plectin deficiency results in muscular dystrophy with epidermolysis bullosa. *Nat Genet*. 1996;13(4):450-457. https://doi.org/10.1038/ng0896-450.
47. Pfendner E, Uitto J. Plectin gene mutations can cause epidermolysis bullosa with pyloric atresia. *J Invest Dermatol*. 2005;124(1):111-115. https://doi.org/10.1111/j.0022-202X.2004.23564.x.
48. Argyropoulou Z, Liu L, Ozoemena L, et al. A novel PLEC nonsense homozygous mutation (c.7159G>T; p.Glu2387*) causes epidermolysis bullosa simplex with muscular dystrophy and diffuse alopecia: a case report. *BMC Dermatol*. 2018;18(1):1-5.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Jacinto JGP, Häfliger IM, Veiga IMB, Drögemüller C, Agerholm JS. A de novo mutation in *KRT5* in a crossbred calf with epidermolysis bullosa simplex. *J Vet Intern Med*. 2020;34:2800–2807. https://doi.org/10.1111/jvim.15943

Clinicopathological and genomic characterization of a Simmental calf with generalized bovine juvenile angiomatosis

Journal:	animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.3390/ani11030624

animals

MDPI

Article

Clinicopathological and Genomic Characterization of a Simmental Calf with Generalized Bovine Juvenile Angiomatosis

Joana G. P. Jacinto [1,2], Irene M. Häfliger [2], Nicole Borel [3], Patrik Zanolari [4], Cord Drögemüller [2,*] and Inês M. B. Veiga [5]

1 Department of Veterinary Medical Sciences, University of Bologna, 40064 Ozzano Emilia (BO), Italy; joana.goncalves2@studio.unibo.it
2 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch
3 Institute of Veterinary Pathology, Vetsuisse Faculty, University of Zurich, 8057 Zurich, Switzerland; n.borel@access.uzh.ch
4 Clinic for Ruminants, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; patrik.zanolari@vetsuisse.unibe.ch
5 Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; ines.veiga@vetsuisse.unibe.ch
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-31-631-2529

Simple Summary: Vascular anomalies represent a heterogeneous group of rare disorders encompassing both vascular malformations and tumors, which can be congenital or arise shortly after birth. They often pose a diagnostic challenge in human and veterinary medicine, and the referring nomenclature is equivocal. Bovine juvenile angiomatosis (BJA), a clinical condition belonging to this group of disorders, encompasses vascular malformations and tumors arising in calves. Usually, such vascular anomalies are not further investigated on a molecular genetic level, mainly because of a lack of resources and diagnostic tools, as well as the low value and short lifespan of the affected animals. Here we report the clinical, pathological, immunohistochemical, and genetic features of a Simmental calf that displayed multiple cutaneous, subcutaneous, and visceral vascular hamartomas compatible with a generalized form of BJA. Whole-genome sequencing identified six coding variants, including four heterozygous variants in the *PREX1, UBE3B, PCDHGA2*, and *ZSWIM6* genes, which occurred only in the BJA-affected calf and were absent in the global control cohort of more than 4500 cattle. Assuming a germline mutation as etiology, one of these variants might be responsible for the vascular malformations identified in this calf.

Abstract: Bovine juvenile angiomatosis (BJA) comprises a group of single or multiple proliferative vascular anomalies in the skin and viscera of affected calves. The purpose of this study was to characterize the clinicopathological phenotype of a 1.5-month-old Simmental calf with multiple cutaneous, subcutaneous, and visceral vascular hamartomas, which were compatible with a generalized form of BJA, and to identify genetic cause for this phenotype by whole-genome sequencing (WGS). The calf was referred to the clinics as a result of its failure to thrive and the presence of multiple cutaneous and subcutaneous nodules, some of which bled abundantly following spontaneous rupture. Gross pathology revealed similar lesions at the inner thoracic wall, diaphragm, mediastinum, pericardium, inner abdominal wall, and mesentery. Histologically, variably sized cavities lined by a single layer of plump cells and supported by a loose stroma with occasional acute hemorrhage were observed. Determined by immunochemistry, the plump cells lining the cavities displayed a strong cytoplasmic signal for PECAM-1, von Willebrand factor, and vimentin. WGS revealed six private protein-changing variants affecting different genes present in the calf and absent in more than 4500 control genomes. Assuming a spontaneous de novo mutation event, one of the identified variants found in the *PREX1, UBE3B, PCDHGA2*, and *ZSWIM6* genes may represent a possible candidate pathogenic variant for this rare form of vascular malformation.

check for updates

Citation: Jacinto, J.G.P.; Häfliger, I.M.; Borel, N.; Zanolari, P.; Drögemüller, C.; Veiga, I.M.B. Clinicopathological and Genomic Characterization of a Simmental Calf with Generalized Bovine Juvenile Angiomatosis. *Animals* **2021**, *11*, 624. https://doi.org/10.3390/ani11030624

Academic Editors: Francesca Ciotola, Sara Albarella and Vincenzo Peretti

Received: 28 January 2021
Accepted: 22 February 2021
Published: 26 February 2021

Keywords: cattle; bovine juvenile angiomatosis; vascular hamartoma; precision medicine; vascular malformation; rare diseases; whole-genome sequencing

1. Introduction

Congenital vascular tumors and malformations are rare anomalies that develop during pregnancy or within the first three months of life [1]. The International Society for the Study of Vascular Anomalies (ISSVA) published an updated classification of these lesions in 2020 that includes genetic and extended histologic findings that came to light since the original ISSVA classification was created in 1996 [2,3]. However, the differentiation between vascular tumors and malformations is not always straightforward [4], and the terminology of vascular anomalies often remains a challenge [2], not only in human, but also in veterinary medicine [5–7]. Specifically, calves are known to display several types of vascular anomalies, some of which are congenital [6]. Vascular hamartomas are relatively common and might be found in the mandibular gingiva [7–11], skin [12], heart [13], and lung [14]. Hemangioma is the most frequently reported benign vascular neoplasm in calves [15] and might be localized in the gingiva [15,16], skin [5,17,18], heart [6], or multifocally [17]. In addition, although malignant vascular tumors are rarely described in cattle [18,19], a multifocal hemangiosarcoma was diagnosed in a stillborn calf [20]. In 1990, Watson and Thompson [6] proposed that these different manifestations of single and multiple vascular anomalies in calves should be grouped under the term bovine juvenile angiomatosis (BJA). This condition differs from the so-called bovine cutaneous angiomatosis, which is mostly identified in adult dairy cattle with a mean age of 5.5 years and is characterized by the appearance of mostly single, proliferative cutaneous vascular anomalies that range from hamartomas to hemangiomas [21–25]. It was postulated that bovine cutaneous angiomatosis may be a consequence of exuberant granulation tissue formation, especially due to histologic similarities to lobular capillary hemangioma or granulation tissue-type hemangioma of man [6,21]. Some authors mentioned chromosomal abnormalities as a putative cause for the bovine juvenile angiomatosis in comparison to what has been described in humans with similar vascular lesions [6,26]. However, chromosomal abnormalities or other types of genetic mutations have not been identified in calves affected by BJA [6,26].

Herein, we aimed to characterize the clinical and pathological phenotype of multiple cutaneous, subcutaneous, and visceral vascular hamartomas compatible with a generalized form of BJA in a Simmental calf. In addition, whole-genome sequencing (WGS) was carried out to identify putatively pathogenic variants.

2. Materials and Methods

2.1. Clinical and Pathological Investigation

A 1.5-month-old female Simmental calf with a body weight of 64 kg was referred to the Clinic for Ruminants at the Vetsuisse Faculty, University of Bern, as a result of poor weight gain and skin lesions. Previous treatment by the referring veterinarian included anthelmintics (Ivermectin 0.2 mg/kg, sc, Ivomec, Biokema SA, Crissier, Switzerland) and antibiotics (Benzylpenicillinum procainum, 30,000 IU/kg, iv, Cobiotic, Virbac AG, Glattburg, Switzerland). The calf was euthanized with an intravenous injection of pentobarbital (pentobarbitalum natricum, 150 mg/kg, iv, Streuli Pharma AG, Uznach, Switzerland) and was subsequently submitted to the Institute of Animal Pathology at the Vetsuisse Faculty, University of Bern, for necropsy and histologic examination. Tissue samples from the subcutaneous and internal nodules, as well as from several inner organs, were immediately collected, fixed in 4% buffered formalin, embedded in paraffin, cut at 4 µm, and stained with haematoxylin and eosin (H&E) for further histologic evaluation. Immunohistochemical (IHC) analysis for platelet endothelial cell adhesion molecule (PECAM-1), von Willebrand factor, smooth muscle actin (SMA), vimentin, and a broad-spectrum cytokeratin

marker (MNF116) were performed from one subcutaneous and one mediastinal nodule. For the PECAM-1 IHC, antigen retrieval using pressure cooking (98 °C, 20 min) in basic EDTA buffer (pH 9.0) was performed, and the primary antibody (sc1506, Santa Cruz Biotechnology, Dallas, TX, USA) was incubated for 1 h at room temperature (RT) (1:1000 dilution). For the von Willebrand factor IHC, pressure cooking in citate buffer (pH 6.0, S2031 Agilent Technologies, Santa Clara, CA, USA) was performed for antigen retrieval, and the primary antibody (A0082, Agilent Technologies) was incubated for 40 min (1:100 dilution) at RT. Peroxidase blocking (S2023, Agilent Technologies) was performed prior to primary antibody incubation for 10 min in both cases, followed by incubation with Envision+system HRP rabbit (K4003, Agilent Technologies) for 30 min at RT, labeling with 3,3′-diaminobenzidine (DAB) (K3468, Agilent Technologies) for 10 min, counterstaining with hematoxylin, and mounting. For the SMA IHC, no antigen retrieval was performed, primary antibody incubation (M085, Agilent Technologies) took place for 1 h at RT (1:400 dilution), and the Mach 4 Universal HRP Polymer kit (BRR 4012L, Medite, Dietikon, Switzerland) was used as secondary antibody at RT. Peroxidase blocking and DAB labeling were performed as previously described. For the vimentin IHC, antigen retrieval was performed with Bond Epitope Retrieval Buffer Type 2 (Tris-EDTA pH 9, AR9640 Leica Biosystems, Wetzlar, Germany) for 10 min at 95 °C, and was followed by primary antibody incubation (V9, M0725, Agilent Technologies) for 15 min at RT (1:1000 dilution). For the MNF116 IHC, antigen retrieval was performed with Bond Enzyme 1 (AR9551 Leica Biosystems) for 5 min at 37 °C, followed by primary antibody incubation (M0821, Agilent Technologies) for 15 min at RT (1:400 dilution). In both cases, all further steps were performed using reagents of the Bond Polymer Refine Detection Kit (DS9800, Leica Biosystems), namely peroxidase blocking for 5 min, incubation with rabbit anti-mouse secondary antibody for 8 min at RT, and peroxidase-labelled polymer incubation for 8 min. Finally, slides were developed in DAB/H_2O_2 for 10 min, counterstained with hematoxylin, and mounted. Tissue sections from normal haired skin from a cow were used as positive control for the PECAM-1, von Willenbrand, vimentin, and MNF116 IHCs, while a tissue section from the small intestine from a cow was used as positive control for the SMA IHC. Tissue sections from the nodules from the affected calf without primary antibody incubation were used as negative controls.

2.2. DNA Sample and Whole-Genome Sequencing

Genomic DNA was isolated from the liver of the calf using the Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland). WGS using the Illumina NovaSeq6000 was performed on the genomic DNA of the calf. The sequenced reads were mapped to the ARS-UCD1.2 reference genome, resulting in an average read depth of approximately 17.5× [27], and single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format files were in accordance with the 1000 Bull Genomes Project processing guidelines of run 8 (www.1000bullgenomes.com) [28], except for the trimming, which was performed using fastp [29]. Further preparation of the genomic data had been done according to Häfliger et al. 2020 [30]. In order to find private variants, we compared the genotypes of the affected calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies at the Institute of Genetics at the Vetsuisse Faculty, University of Bern, and that are publicly available in the European Nucleotide Archive (SAMEA6528880 is the sample accession number of the affected calf; http://www.ebi.ac.uk/en) (Table S1). The filtered list of remaining variants were further checked for their occurrence in a global control cohort of 4110 genomes of a variety of breeds (1000 Bull Genomes Project run 8; www.1000bullgenomes.com accessed on 15 November 2020) [28]. Integrative Genomics Viewer (IGV) [31] software was used for visual inspection of genome regions containing possible candidate genes.

In order to evaluate possible chromosomal abnormalities, the read depth along all chromosomes was calculated. A sliding window approach was used where 3 different window sizes were executed (10 kb, 200 kb, 500 kb). Using the function bedcov of the

program Samtools [32], the output generated was the number of reads within each specified window. Furthermore, coverage plots were produced using the function Manhattan of the package "qqman" in R [33].

2.3. Evaluation of the Molecular Consequences of Amino Acid Substitutions

PROVEAN [34], MutPred2 [35], and PredictSNP1 [36] were used to predict the functional consequences of the identified variants on protein.

3. Results

3.1. Clinical Phenotype

At clinical examination, the calf was alert but moderately reduced in its general body condition. The rectal body temperature was 39.0 °C, the heart rate was 80 beats per minute, and the respiratory rate was 40 beats per minute. Examination of the cardiovascular, respiratory, digestive, and urinary systems did not reveal any abnormalities.

The integumentary system of the calf revealed multiple, cutaneous and subcutaneous, soft, movable nodules, which measured up to 5 cm in diameter. These were predominantly present in the caudal aspect of the back, the pelvis, and the hind limbs (Figure 1a). The cutaneous nodules were occasionally covered by sanguineous crusts, and some nodules bled abundantly following spontaneous rupture (Figure 1b). No hematological and biochemical analyses were performed due to cost restrictions. Based on these findings and considering that no other calves in the herd showed similar skin lesions, a presumptive diagnosis of multifocal vascular anomalies was made. The animal was then euthanized due to a lack of response to treatment and poor prognosis.

Figure 1. Cutaneous and subcutaneous nodules (arrows) located at the back, flank, and hindlimbs

of the Simmental calf (**a**), some of which were covered by abundant serosanguineous crusts (**b**). At necropsy, similar looking nodules (arrows) were identified, namely in the inner thoracic wall (**c**) and in the mediastinum (**d**). All nodules measured up to 5 cm in diameter, were well demarcated from the adjacent tissue, and displayed a white to reddish cut surface. Bar 5 cm (**b**), 3 cm (**c**), and 4 cm (**d**).

3.2. Pathological Phenotype

At necropsy, the cutaneous and subcutaneous nodules displayed either a white or reddish cut surface. Additionally, mostly pedunculated but similar looking nodules were occasionally present at the inner thoracic wall and diaphragm (Figure 1c), the mediastinum (Figure 1d), the pericardium, the inner abdominal wall, and the mesentery adjacent to the duodenum. The remaining organs were macroscopically unremarkable.

Histologically, all analyzed nodules consisted of encapsulated (with the exception of the nodule present in the mesentery adjacent to the duodenum), well demarcated, expansively growing, moderately cellular masses. These consisted of abundant, variably sized, mostly empty cavities, which were lined by a single layer of plump cells and supported by a loose, partially edematous fibrovascular stroma (Figure 2a,b), with occasionally intermingled foci of mature connective tissue. The plump cells displayed a moderate amount of eosinophilic, homogeneous cytoplasm, an oval nucleus with finely stippled chromatin, and up to two basophilic, round nucleoli. The anisocytosis and anisokaryosis were low to moderate, and there were very few mitotic figures visible. Moderate to high numbers of free erythrocytes (compatible with acute hemorrhage), as well as occasional neutrophils and lymphocytes, were visible within the stroma.

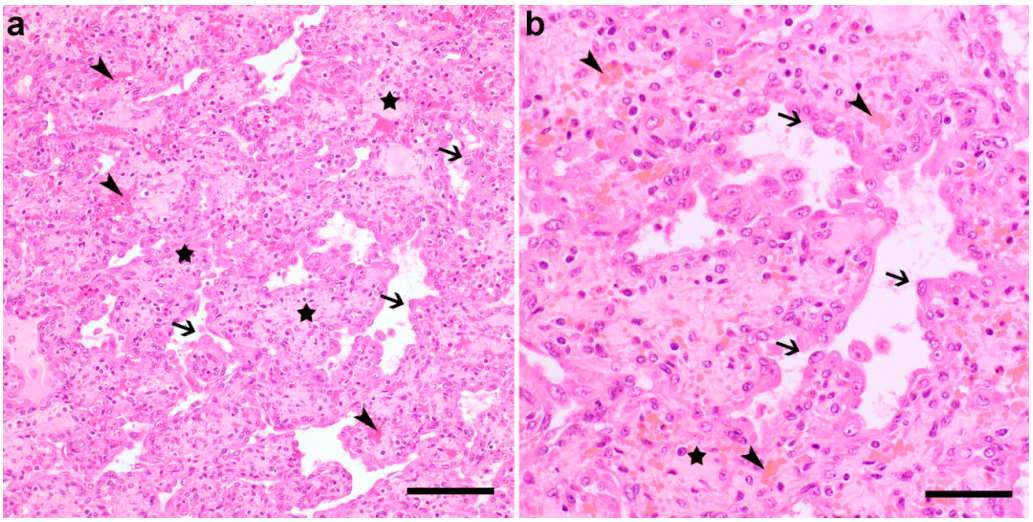


Figure 2. Histology of one of the subcutaneous nodules from the Bovine juvenile angiomatosis (BJA)-affected Simmental calf. Note the variably sized, mostly empty cavities lined by a single layer of plump cells (arrows) and supported by a loose fibrovascular stroma (stars) with occasional acute hemorrhage (arrowheads). Hematoxylin and eosin (H&E) staining, Bar 100 µm (**a**) and 50 µm (**b**).

To determine the cellular origin of these nodules, IHC was performed. In both nodules, the plump cells lining the cavities displayed a strong cytoplasmic signal for PECAM-1 (Figure 3a), von Willebrand factor (Figure 3b), and vimentin (Figure 3c) but not for SMA (Figure 3d). In addition, the supporting stromal cells displayed a strong cytoplasmic signal for vimentin (Figure 3c) and SMA (Figure 3d), as well as a rather strong background staining in the von Willebrand factor IHC (Figure 3b). Cells lining the cavities and stromal cells were negative for the pan cytokeratin marker MNF116 (not shown).

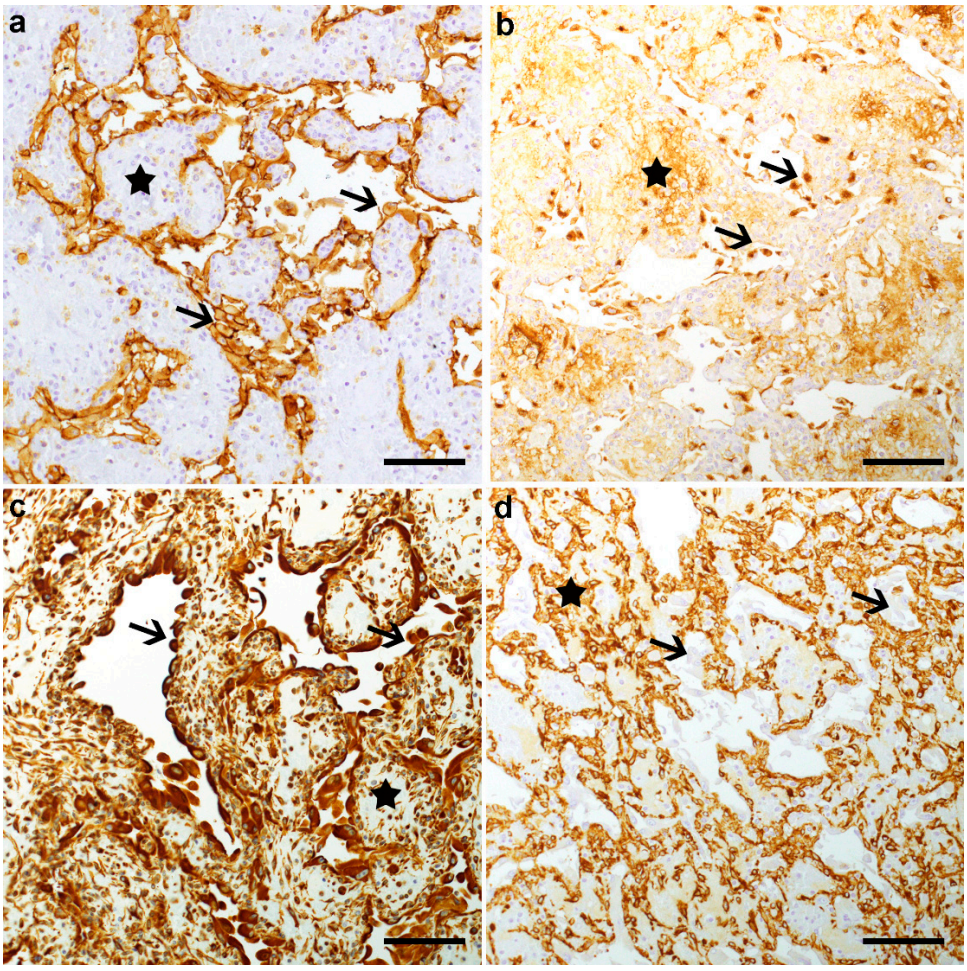


Figure 3. Immunohistochemical (IHC) analysis of a subcutaneous nodule from the BJA-affected Simmental calf with platelet endothelial cell adhesion molecule (PECAM-1) (**a**), von Willebrand factor (**b**), vimentin (**c**), and smooth muscle actin (SMA) (**d**). Note that the plump cells lining the cavities (arrows) displayed a strong positive signal in the PECAM-1, von Willebrand factor, and vimentin IHC but were negative in the SMA IHC. The stromal cells (stars) displayed a positive cytoplasmic signal in the vimentin and SMA IHC staining, while the signal seen in the von Willebrand factor IHC was considered to represent background staining. IHC staining, Hematoxylin counterstain, Bar 100 µm.

The above described histologic and immunohistochemical features resemble previously described vascular anomalies in calves [7,8]. In spite of the often conflicting nomenclature terminology [6,7], a final diagnosis of multiple cutaneous, subcutaneous, and visceral vascular hamartomas compatible with a generalized form of BJA as described by Watson and Thompson [6] was made in this case.

3.3. Genetic Analysis

Assuming a spontaneous mutation as etiology for this most likely congenital condition, the sequencing of the whole genome of the affected calf was carried out. Filtering of the obtained variant catalogue for private variants exclusively present in the BJA-affected calf and absent in 496 available control genomes identified 31 private protein-changing variants, and subsequent visual inspection using IGV software confirmed 29 as real variants. Analyzing the occurrence of these variants in the global control cohort of 4110 genomes of a variety of breeds [28], six heterozygous protein-changing variants exclusively present in the genome of the affected calf were identified (Table S2). A total of 137 sequenced genomes from Simmental cattle were considered during variant filtering. Results of the functional impact prediction of these six heterozygous protein-changing variants are presented in Table 1.

Table 1. Pathogenicity prediction results for the six heterozygous protein-changing variants exclusively present in the genome of the BJA-affected calf and absent in global control cohort of more than 4500 genomes of a variety of breeds.

Gene	Effect	Protein-Changing	pLI [1]	PROVEAN Score	PROVEAN Impact	MutPred2 Score	MutPred2 Impact	PredictSNP1 Score	PredictSNP1 Impact
PREX1	missense	p.Arg401Cys	1	−5.149	deleterious	0.387	neutral	0.719	deleterious
UBE3B	missense	p.Ala32Val	0	−1.946	neutral	0.583	neutral	0.510	deleterious
PCDHGA2	disruptive in-frame deletion	p.Lys141_Val142del	0	−11.366	deleterious	0.338	neutral	NA	NA
ZSWIM6	disruptive in-frame deletion	p.Ala146_Gly148del	1	1.280	neutral	0.366	neutral	NA	NA
NR1H3	missense	p.Thr46Met	0.9	−0.371	neutral	0.086	neutral	0.653	neutral
C23H6orf132	missense	p.Gly692Glu	NA	−0.907	neutral	0.035	neutral	0.826	neutral

[1] probability of loss-of-function intolerance score (pLI) provided by the Genome Aggregation Database (gnomAD) [37]. NA, not available.

Based on the function of the identified genes and in the predicted impact on the protein, the variants identified in *PREX1*, *UBE3B*, *ZSWIM6*, and *PCDHGA2* were considered the most likely candidate pathogenic variants for the observed phenotype. Unfortunately, no biological samples of the sire and dam were available to evaluate if one of these variants has occurred de novo in the calf.

No evidence for chromosomal abnormalities were detected by analyzing the obtained read depth or coverage along all chromosomes.

4. Discussion

We performed a comprehensive clinical, pathological, and genetic investigation in a Simmental calf displaying multifocal subcutaneous and visceral vascular hamartomas compatible with a generalized form of BJA.

We evaluated the hypothesis of a spontaneous mutation as the possible cause for this most likely congenital phenotype. Analysis of the genome sequence revealed six heterozygous protein-changing variants in the *PREX1*, *UBE3B*, *PCDHGA2*, *ZSWIM6*, *NR1H3*, and *C23H6orf132* genes that were exclusively present in the genome of the affected calf and absent in a global control cohort of more than 4500 cattle genomes of a variety of breeds. Therefore, we considered these apparently rare coding variants as possible candidates for the observed BJA phenotype. In the following, literature and in silico effect predictions are used to discuss a conceivable causal role. The missense variants identified in the *NR1H3* and *C23H6orf132* genes were not considered as putative causes for these vascular malformations because they were predicted to have a neutral or benign effect.

However, the missense variant found in *PREX1* was predicted to be deleterious by different tools such as PROVEAN and PredictSNP1 [34,36]. This gene belongs to the family of Rac guanine nucleotide exchange factors (Rac-GEF) and is activated by phosphatidylinositol 3,4,5-trisphosphate (PI (3,4,5) P3), which is generated by class I phosphoinositide 3-kinase (*PI3K*) and the β-gamma subunits of the heterotrimeric-G proteins (Gβγ) [38]. Furthermore, *PREX1* has an important role in the control of many fundamental cellular functions, including cell migration, actin cytoskeletal rearrangement, adhesion, and the production of reactive oxygen species (ROS) [39]. Additionally, the major effector of PREX1 protein activity is related to the induction of actin-mediated membrane ruffling and lamellipodia production at the leading edge of cell migration, and abnormally activated Rac is involved in the metastasis and invasion of tumor cells [40]. Evidence suggested that Rac and PREX1 protein are increased in cell proliferation and migration in several human cancers such as melanoma [41], breast cancer [42], prostate cancer [43], and oral squamous cell carcinoma [44]. However, no association between *PREX1* and the occurrence of vascular tumors has been reported to date.

The missense variant in *UBE3B* was predicted to be deleterious by PredictSNP1 [36]. An independent splice site variant in bovine *UBE3B* is associated with an autosomal-recessive inherited disorder called PIRM syndrome in Finnish Ayrshire cattle, which causes

intellectual disability, retarded growth, and mortality (OMIA 001934-9913) [45], and which resembles the human autosomal-recessive Kaufman oculocerebrofacial syndrome (OMIM 244450) [46]. In addition, a recent study demonstrated that suppression of the E3 ubiquitin ligase UBE3B-mediated MYC ubiquitination and degradation caused by the integration of *TRIB3* with *MYC* is associated with high proliferation and self-renewal of lymphoma cells [47]. However, in the case presented in this study, the identified variant in *UBE3B* was heterozygous, and the calf did not show a phenotype compatible with these disorders.

Moreover, a disruptive in-frame deletion in *PCDHGA2* was predicted to have a deleterious impact in the protein using PROVEAN [34]. This gene presents a probability of loss-of-function intolerance score (pLI) of zero according to Genome Aggregation Database (gnomAD) [37]. Considering that transcripts with a pLI superior or equal to 0.9 are predicted to be loss-of-function (LoF) intolerant due to haploinsufficiency of the gene [48], *PCDHGA2* most likely does not belong to the group of LoF haploinsufficient genes. In addition, PCDHGA2 protein is a potential calcium-dependent cell-adhesion protein that might be involved in the establishment and maintenance of specific neuronal connections in the brain [49,50]. In humans, somatic mutations in *PCDHGA2* have been associated with cell and biological adhesion in aggressive papillary thyroid microcarcinomas [51]. However, and similarly to *PREX1*, no association between this gene and the occurrence of vascular tumors has been reported to date.

To date, no pathogenic variants of *ZSWIM6* have been reported in domestic animals. Interestingly, recent data obtained from human genome sequencing studies presented in the gnomAD [37] showed that the pLI for this gene was 1, meaning that *ZSWIM6* falls into the class of LoF haploinsufficient genes. The encoded protein, so-called Zinc finger SWIM domain-containing protein 6, is a protein of unknown function that is involved in the nervous system development and regulation [52]. Moreover, *ZSWIM6* enables the zinc ion binding [53] and is part of the Cul2-RING ubiquitin ligase complex [49]. In humans, pathogenic variants in the *ZSWIM6* gene (OMIM 615951) are associated to acromelic frontonasal dysostosis [54] and neurodevelopmental disorder with movement abnormalities, abnormal gait, and autistic features [52,55]. Additionally, the ZSWIM6 protein, according to the Biological General Repository for Interaction Datasets (BioGRID), is predicted to interact physically with ARIH1 [56], GLMN [57], HECW2 [58], HNRNPH1 [59] and HNRNPL [60]. Similarly to ZSWIM6, the GLMN is part of the Cul2-RING ubiquitin ligase complex [61] and known to be involved in numerous different processes, namely the normal development of the vasculature [62]. Particularly, autosomal dominant pathogenic variants in the *GLMN* gene (OMIM 138000) are associated with the development of glomuvenous malformations (GVMs) [63] and Blue rubber bleb nevus syndrome (BRBNS) (OMIM 112200) [64] in humans.

Based on these findings, we speculate whether the above-described vascular lesions in this calf could be due to an impaired interaction between ZSWIM6 and GLMN. However, it cannot be excluded that the phenotype displayed by this calf is due to one of the other remaining variants found exclusively in the sequenced case, or even to a combination of these. Either way, we hypothesize that one of these variants either occurred post-zygotically during the fetal development of the affected calf or represent a germline mutation that occurred in one of the parents. To prove that the private variants occurred was indeed de novo, genotyping of the sire and dam would be needed. Unfortunately, no samples from these animals were available at the time the genetic analysis was performed, therefore this hypothesis cannot be confirmed.

We hope that these findings may contribute to a better knowledge and characterization of BJA. However, it is highly unlikely that the candidate causal variants identified in the genome of the studied calf are responsible for other BJA cases since this condition encompasses several kinds of vascular anomalies, including both vascular malformations and tumors [6].

5. Conclusions

This report highlights the utility of WGS-based precision diagnostics for understanding the underlying genetics of rare disorders in animals with an available reference genome sequence and the value of surveillance for harmful genetic disorders in cattle breeding populations.

Supplementary Materials: The following are available online at https://www.mdpi.com/2076-2615/11/3/624/s1, Table S1: Project and Sample ID for public access of the whole genome sequenced genomes used in this study. EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available, Table S2: List of the remaining variants after the comparison to the global control cohort of 4110 genomes of other breeds (1000 Bull Genomes Project run 8; www.1000bullgenomes.com), revealing 6 protein-changing variants with a predicted moderate impact only present in the affected calf.

Author Contributions: Conceptualization, C.D. and I.M.B.V.; methodology, I.M.H., N.B., P.Z. and I.M.B.V.; software, I.M.H.; validation, J.G.P.J., N.B. and C.D.; formal analysis, J.G.P.J. and I.M.H.; investigation, J.G.P.J., N.B. and I.M.B.V.; resources, N.B., P.Z., C.D. and I.M.B.V.; data curation, C.D.; writing—original draft preparation, J.G.P.J. and I.M.B.V.; writing—review and editing, J.G.P.J., C.D., N.B. and I.M.B.V.; visualization, J.G.P.J., N.B. and I.M.B.V.; supervision, C.D.; project administration, I.M.B.V. and C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by the Swiss National Science Foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but part of a clinical and pathological veterinary diagnostic case. The animal in this study was examined and euthanized with the consent of its owner.

Data Availability Statement: The whole-genome data of the affected calf is freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA6528880.

Acknowledgments: We thank Nathalie Besuchet Schmutz, Hans Gantenbein, Manuela Bozzo, Erika Bürgi, and Bettina de Breuyn for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alamo, L.; Beck-Popovic, M.; Gudinchet, F.; Meuli, R. Congenital tumors: Imaging when life just begins. *Insights Imaging* **2011**, *2*, 297–308. [CrossRef]
2. Wassef, M.; Blei, F.; Adams, D.; Alomari, A.; Baselga, E. Vascular Anomalies Classi fi cation: Recommendations From the International Society for the Study of Vascular Anomalies. *Pediatrics* **2015**, *136*, 203–214. [CrossRef] [PubMed]
3. ISSVA Classification of Vascular Anomalies. Available online: issva.org/classification (accessed on 23 November 2020).
4. Mansfield, S.A.; Williams, R.F.; Iacobas, I. Vascular tumors. *Semin. Pediatr. Surg.* **2020**, *29*. [CrossRef]
5. Priestnall, S.L.; De Bellis, F.; Bond, R.; Alony-Gilboa, Y.; Summers, B.A. Spontaneous regression of congenital cutaneous hemangiomas in a calf. *Vet. Pathol.* **2010**, *47*, 343–345. [CrossRef]
6. Watson, T.D.; Thompson, H. Juvenile bovine angiomatosis: A syndrome of young cattle. *Vet. Rec.* **1990**, *127*, 279–282.
7. Rösti, L.; Lauper, J.; Merhof, K.; Gorgas, D.; Ross, S.; Grest, P.; Welle, M. Angeborene Gefässanomalien in der Maulhöhle bei zwei Kälbern. *Schweiz Arch Tierheilkd* **2013**, 627–632. [CrossRef]
8. Sheahan, B.J.; Donnelly, W. Vascular hamartoma in the gingiva of two neonatal calves. *J. Am. Vet. Med. Assoc.* **1981**, *184*, 205–206.
9. Yeruham, I.; Abramovitch, I.; Perl, S. Gingival vascular hamartoma in two calves. *Aust. Vet. J.* **2004**, *82*, 152–153. [CrossRef]
10. Stanton, M.; Meunier, P.; Smith, D. Vascular hamartoma in the gingiva of two neonatal calves. *J. Am. Vet. Med. Assoc.* **1984**, *184*, 205–206.
11. Wilson, R.B. Gingival vascular hamartoma in three calves. *J. Vet. Diagnostic Investig.* **1990**, *2*, 338–339. [CrossRef]
12. Yeruham, I.; Perl, S.; Orgad, U. Congenital skin neoplasia in cattle. *Vet. Dermatol.* **1999**, *10*, 149–156. [CrossRef]
13. Brisville, A.C.; Buczinski, S.; Chénier, S.; Francoz, D. A cardiac vascular hamartoma in a calf: Ultrasonographic and pathologic images. *J. Vet. Cardiol.* **2012**, *14*, 377–380. [CrossRef]
14. Roth, L.; Bradley, G.A. Pulmonary hamartoma in a calf. *J. Comp. Pathol.* **1991**, *105*, 471–474. [CrossRef]
15. Misdorp, W. Tumours in calves: Comparative aspects. *J. Comp. Pathol.* **2002**, *127*, 96–105. [CrossRef] [PubMed]

16. Tontis, A. Kongenitales kavernöses Hämangiom beim Brown-Swiss-Kalb, ein seletenes orales Blastom. *Tierarztl. Prax.* **1994**, *22*, 137–139.
17. Baker, J.C.; Hultgren, B.D.; Larson, V.L. Disseminated cavernous hemangioma in a calf. *J. Am. Vet. Med. Assoc.* **1982**, *181*, 172–173.
18. Poulsen, K.; McSloy, A.; Perrier, M.; Prichard, M.; Steinberg, H.; Semrad, S. Primary mandibular hemangiosarcoma in a bull. *Can. Vet. J.* **2008**, *49*, 901–903. [PubMed]
19. Stock, M.L.; Smith, B.I.; Engiles, J.B. Disseminated hemangiosarcoma in a cow. *Can. Vet. J.* **2011**, *52*, 409–413.
20. Badylak, S.F. Congenital Multifocal Hemangiosarcoma in a Stillborn Calf. *Vet. Pathol.* **1983**, *20*, 245–247. [CrossRef]
21. Hendrick, M.J. Mesenchymal Tumors of the Skin and Soft Tissues. *Tumors Domest. Anim.* **2016**, 142–175. [CrossRef]
22. Cotchin, E.; Swarbrick, O. Bovine cutaneous angiomatosis: A lesion resembling human pyogenic granuloma (granuloma telangiecticum). *Vet. Rec.* **1963**, *75*, 437–444.
23. Lombard, C.; Levesque, L. H'emangiomatose hyperplasique cutan'ee de vaches Normandes. *Ann. Anat. Pathol. (Paris)* **1964**, *9*, 453–462. [PubMed]
24. Moulton, J.E. *Tumors in Domestic Animals*, 2nd ed.; University of California Press: Berkeley, CA, USA, 1978.
25. Waldvogel, A.; Hauser, B. "Bovine kutane Angiomatose" in der Schweiz. *Schweiz. Arch. Tierheilkd.* **1983**, *125*, 329–333. [PubMed]
26. Schilcher, F.; Gasteiner, J.; Palmetshofer, G.; Baumgartnerm, W. Generalisierte juvenile Angiomatose bei einem Kalb. *Wien. Tierarztl. Monatsschr.* **2000**, *87*, 10–13.
27. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, 1–9. [CrossRef]
28. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef] [PubMed]
29. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef] [PubMed]
30. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
31. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant review with the integrative genomics viewer. *Cancer Res.* **2017**, *77*, e31–e34. [CrossRef]
32. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [CrossRef]
33. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2013.
34. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]
35. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.; Lin, G.N.; Nam, H.-J.; Mort, M.; Cooper, D.; Sebat, J.; Iakoucheva, L.; et al. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *Nat. Commun.* **2017**, 134981. [CrossRef]
36. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Comput. Biol.* **2014**, *10*, 1–11. [CrossRef] [PubMed]
37. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]
38. Welch, H.C.E.; Coadwell, W.J.; Ellson, C.D.; Ferguson, G.J.; Andrews, S.R.; Erdjument-Bromage, H.; Tempst, P.; Hawkins, P.T.; Stephens, L.R. P-Rex1, a PtdIns(3,4,5)P3- and gβγ-regulated guanine-nucleotide exchange factor for Rac. *Cell* **2002**, *108*, 809–821. [CrossRef]
39. Jaffe, A.B.; Hall, A. Rho GTPases: Biochemistry and biology. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 247–269. [CrossRef] [PubMed]
40. Vigil, D.; Cherfils, J.; Rossman, K.L.; Der, C.J. Ras superfamily GEFs and GAPs: Validated and tractable targets for cancer therapy? *Nat. Rev. Cancer* **2010**, *10*, 842–857. [CrossRef] [PubMed]
41. Ryan, M.B.; Finn, A.J.; Pedone, K.H.; Thomas, N.E.; Der, C.J.; Cox, A.D. ERK/MAPK signaling drives overexpression of the rac-GEF, PREX1, in BRAF-and NRAS-mutant melanoma. *Mol. Cancer Res.* **2016**, *14*, 1009–1018. [CrossRef] [PubMed]
42. Sosa, M.S.; Lopez-Haber, C.; Yang, C.; Wang, H.B.; Lemmon, M.A.; Busillo, J.M.; Luo, J.; Benovic, J.L.; Klein-Szanto, A.; Yagi, H.; et al. Identification of the Rac-GEF P-Rex1 as an Essential Mediator of ErbB Signaling in Breast Cancer. *Mol. Cell* **2010**, *40*, 877–892. [CrossRef]
43. Goel, H.L.; Pursell, B.; Shultz, L.D.; Greiner, D.L.; Brekken, R.A.; Vander Kooi, C.W.; Mercurio, A.M. P-Rex1 Promotes Resistance to VEGF/VEGFR-Targeted Therapy in Prostate Cancer. *Cell Rep.* **2016**, *14*, 2193–2208. [CrossRef] [PubMed]
44. Wan, S.C.; Wu, H.; Li, H.; Deng, W.W.; Xiao, Y.; Wu, C.C.; Yang, L.L.; Zhang, W.F.; Sun, Z.J. Overexpression of PREX1 in oral squamous cell carcinoma indicates poor prognosis. *J. Mol. Histol.* **2020**, *51*, 531–540. [CrossRef]
45. Venhoranta, H.; Pausch, H.; Flisikowski, K.; Wurmser, C.; Taponen, J.; Rautala, H.; Kind, A.; Schnieke, A.; Fries, R.; Lohi, H.; et al. In frame exon skipping in UBE3B is associated with developmental disorders and increased mortality in cattle. *BMC Genom.* **2014**, *15*, 1–9. [CrossRef]

46. Flex, E.; Ciolfi, A.; Caputo, V.; Fodale, V.; Leoni, C.; Melis, D.; Bedeschi, M.F.; Mazzanti, L.; Pizzuti, A.; Tartaglia, M.; et al. Loss of function of the E3 ubiquitin-protein ligase UBE3B causes kaufman oculocerebrofacial syndrome. *J. Med. Genet.* **2013**, *50*, 493–499. [CrossRef] [PubMed]
47. Li, K.; Wang, F.; Yang, Z.N.; Zhang, T.T.; Yuan, Y.F.; Zhao, C.X.; Yeerjiang, Z.; Cui, B.; Hua, F.; Lv, X.X.; et al. TRIB3 promotes MYC-associated lymphoma development through suppression of UBE3B-mediated MYC degradation. *Nat. Commun.* **2020**, *11*. [CrossRef] [PubMed]
48. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [CrossRef]
49. Gaudet, P.; Livstone, M.S.; Lewis, S.E.; Thomas, P.D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief. Bioinform.* **2011**, *12*, 449–462. [CrossRef] [PubMed]
50. Uniprot UniProtKB - Q9Y5H1. Available online: https://www.uniprot.org/uniprot/Q9Y5H1 (accessed on 16 December 2020).
51. Song, J.; Wu, S.; Xia, X.; Wang, Y.; Fan, Y.; Yang, Z. Cell adhesion-related gene somatic mutations are enriched in aggressive papillary thyroid microcarcinomas 06 Biological Sciences 0604 Genetics. *J. Transl. Med.* **2018**, *16*, 1–9. [CrossRef]
52. Tischfield, D.J.; Saraswat, D.K.; Furash, A.; Fowler, S.C.; Fuccillo, M.V.; Anderson, S.A. Loss of the neurodevelopmental gene Zswim6 alters striatal morphology and motor regulation. *Neurobiol. Dis.* **2017**, *103*, 174–183. [CrossRef]
53. InterPro Zinc finger, SWIM-type (IPR007527). Available online: http://www.ebi.ac.uk/interpro/entry/InterPro/IPR007527/ (accessed on 24 November 2020).
54. Twigg, S.R.F.; Ousager, L.B.; Miller, K.A.; Zhou, Y.; Elalaoui, S.C.; Sefiani, A.; Bak, G.S.; Hove, H.; Hansen, L.K.; Fagerberg, C.R.; et al. Acromelic frontonasal dysostosis and ZSWIM6 mutation: Phenotypic spectrum and mosaicism. *Clin. Genet.* **2016**, *90*, 270–275. [CrossRef]
55. Palmer, E.E.; Kumar, R.; Gordon, C.T.; Shaw, M.; Hubert, L.; Carroll, R.; Rio, M.; Murray, L.; Leffler, M.; Dudding-Byth, T.; et al. A Recurrent De Novo Nonsense Variant in ZSWIM6 Results in Severe Intellectual Disability without Frontonasal or Limb Malformations. *Am. J. Hum. Genet.* **2017**, *101*, 995–1005. [CrossRef]
56. Scott, D.C.; Rhee, D.Y.; Duda, D.M.; Kelsall, I.R.; Olszewski, J.L.; Paulo, J.A.; de Jong, A.; Ovaa, H.; Alpi, A.F.; Harper, J.W.; et al. Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **2016**, *166*, 1198–1214. [CrossRef]
57. Huttlin, E.L.; Bruckner, R.J.; Paulo, J.A.; Cannon, J.R.; Ting, L.; Baltier, K.; Colby, G.; Gebreab, F.; Gygi, M.P.; Parzen, H.; et al. Architecture of the human interactome defines protein communities and disease networks. *Nature* **2017**, *545*, 505–509. [CrossRef] [PubMed]
58. Lu, L.; Hu, S.; Wei, R.; Qiu, X.; Lu, K.; Fu, Y.; Li, H.; Xing, G.; Li, D.; Peng, R.; et al. The HECT type ubiquitin ligase NEDL2 is degraded by anaphase-promoting complex/cyclosome (APC/C)-Cdh1, and its tight regulation maintains the metaphase to anaphase transition. *J. Biol. Chem.* **2013**, *288*, 35637–35650. [CrossRef]
59. Uren, P.J.; Bahrami-Samani, E.; de Araujo, P.R.; Vogel, C.; Qiao, M.; Burns, S.C.; Smith, A.D.; Penalva, L.O.F. High-throughput analyses of hnRNP H1 dissects its multi-functional aspect. *RNA Biol.* **2016**, *13*, 400–411. [CrossRef] [PubMed]
60. Fei, T.; Chen, Y.; Xiao, T.; Li, W.; Cato, L.; Zhang, P.; Cotter, M.B.; Bowden, M.; Lis, R.T.; Zhao, S.G.; et al. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E5207–E5215. [CrossRef]
61. Tron, A.E.; Arai, T.; Duda, D.M.; Kuwabara, H.; Olszewski, J.L.; Fujiwara, Y.; Bahamon, B.N.; Signoretti, S.; Schulman, B.A.; DeCaprio, J.A. The Glomuvenous Malformation Protein Glomulin Binds Rbx1 and Regulates Cullin RING Ligase-Mediated Turnover of Fbw7. *Mol. Cell* **2012**, *46*, 67–78. [CrossRef]
62. Brouillard, P.; Boon, L.M.; Mulliken, J.B.; Enjolras, O.; Ghassibé, M.; Warman, M.L.; Tan, O.T.; Olsen, B.R.; Vikkula, M. Mutations in a novel factor, glomulin, are responsible for glomuvenous malformations ("glomangiomas"). *Am. J. Hum. Genet.* **2002**, *70*, 866–874. [CrossRef] [PubMed]
63. Irrthum, A.; Brouillard, P.; Enjolras, O.; Gibbs, N.F.; Eichenfield, L.F.; Olsen, B.R.; Mulliken, J.B.; Boon, L.M.; Vikkula, M. Linkage disequilibrium narrows locus for venous malformation with glomus cells (VMGLOM) to a single 1.48 Mbp YAC. *Eur. J. Hum. Genet.* **2001**, *9*, 34–38. [CrossRef] [PubMed]
64. Yin, J.; Qin, Z.; Wu, K.; Zhu, Y.; Hu, L.; Kong, X. Rare Germline GLMN Variants Identified from Blue Rubber Bleb Nevus Syndrome Might Impact mTOR Signaling. *Comb. Chem. High Throughput Screen.* **2019**, *22*, 675–682. [CrossRef]

A homozygous missense variant in laminin subunit beta 1 as candidate causal mutation of hemifacial microsomia in Romagnola cattle

Journal:	Journal of Veterinary Internal Medicine
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, performing genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.1111/jvim.16316

Received: 11 June 2021 | Accepted: 5 November 2021

DOI: 10.1111/jvim.16316

CASE REPORT

Journal of Veterinary Internal Medicine Open Access

ACVIM American College of Veterinary Internal Medicine

A homozygous missense variant in laminin subunit beta 1 as candidate causal mutation of hemifacial microsomia in Romagnola cattle

Joana G. P. Jacinto[1,2] | Irene M. Häfliger[2] | Marco Bernardini[3,4] | Maria Teresa Mandara[5] | Ezio Bianchi[6] | Marilena Bolcato[1] | Noemi Romagnoli[1] | Arcangelo Gentile[1] | Cord Drögemüller[2]

[1]Department of Veterinary Medical Sciences, University of Bologna, Ozzano, Italy

[2]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[3]Anicura Portoni Rossi Veterinary Hospital, Zola Predosa, Bologna, Italy

[4]Department of Animal Medicine, Productions and Health, University of Padua, Padua, Italy

[5]Department of Veterinary Medicine, Neuropathology Laboratory, University of Perugia, Perugia, Italy

[6]Department of Veterinary Medical Sciences, University of Parma, Parma, Italy

Correspondence
Joana G. P. Jacinto, Department of Veterinary Medical Sciences, University of Bologna, Ozzano, Italy.
Email: joana.goncalves2@studio.unibo.it

Abstract

Hemifacial microsomia (HFM) was diagnosed in a 9-day-old Romagnola calf. The condition was characterized by microtia of the left ear, anotia of the right ear, asymmetry of the face, and deafness. Magnetic resonance imaging revealed agenesis of the right pinna and both tympanic bullae, asymmetry of the temporal bones and temporomandibular joints, and right pontine meningocele. Brainstem auditory evoked responses confirmed the impaired auditory capacity. At gross post mortem examination, there was agenesis and hypoplasia of the right and the left external ear, respectively. No histological abnormalities were detected in the inner ears. A trio whole-genome sequencing approach was carried out and identified a private homozygous missense variant in *LAMB1* affecting a conserved residue (p.Arg668Cys). Genotyping of 221 Romagnola bulls revealed a carrier prevalence <2%. This represents a report of a *LAMB1*-related autosomal recessive inherited disorder in domestic animals and adds LAMB1 to the candidate genes for HFM.

KEYWORDS

Bos taurus, development, microtia, precision medicine, rare disease, WGS

1 | INTRODUCTION

Microtia is a congenital malformation of the external ear and can range in severity from mild structural abnormalities to complete absence of the ear (anotia).[1] It occurs as an isolated malformation (nonsyndromic form) or as a part of a spectrum of anomalies (syndromic form). Hemifacial microsomia (HFM) is the term used to describe a syndromic form that might be characterized by microtia, facial asymmetry, oral clefts, and eyelid defects. Renal abnormalities, cardiac defects, polydactyly, and vertebral deformities are ancillary malformations.[2,3]

The causes of microtia are poorly understood in both humans and animals,[4] although evidence supports contribution of genetic and environmental components. In humans, there are several monogenic inherited mostly syndromic forms of microtia (OMIM 600674 occur associated with disease-causing variants in genes such as *HOXA1*,[5]

Abbreviations: BAs, brachial arches; BAERs, brainstem auditory evoked responses; CNCCs, cranial neural crest cells; H&E, hematoxylin and eosin; HFM, hemifacial microsomia; IGV, Integrative Genomics Viewer; MRI, magnetic resonance imaging; NCCs, neural crest cells; NHL, normal hearing level; WGS, whole-genome sequencing; EGF, epidermal growth factor.

HOXA2,[6,7] *ORC1*, *ORC6*, *CDT1*,[8] *ORC4*,[9] *CDC6*,[10] *MCM5*,[11] *TCOF1*,[12] *POLR1C*, *POLR1D*,[13] *POLR1B*,[14] and *FGF3*).[15] Mouse model studies identify a list of genes associated with microtia, and illustrated several signaling pathways, including BMP, WNT, FGF, and retinoic acid, that present an important function in outer-ear development.[2] Furthermore, in cattle and sheep dominantly inherited nonsyndromic forms of anomalies affecting the outer ear are associated with regulatory variants affecting the expression of *HMX1*[16,17] (OMIA 000317-9913 and OMIA 001952-9940). A recessive syndromic form of microtia in pigs is associated with a deletion in *HOXA1*[4] (OMIA 001952-9823).

The aims of this study were to describe the clinical and disease phenotype observed in a Romagnola calf affected by HFM, to identify the suspected genetic etiology by a trio-based whole-genome sequencing (WGS) approach, and to estimate the prevalence of the deleterious allele in Romagnola cattle.

2 | CASE DESCRIPTION

A 9-day-old female Romagnola calf, weighting 43 kg, was admitted to the Department of Veterinary Medical Sciences, University of Bologna because absence of the auricles and facial asymmetry.

At the time of admission, the calf had asymmetry of the face with deviation to the right side and lingual ptosis (Figure 1A). The right pinna was absent (anotia), while the left 1 was a rudiment of soft tissue with absence of the ear canal (aural atresia) and covered by long

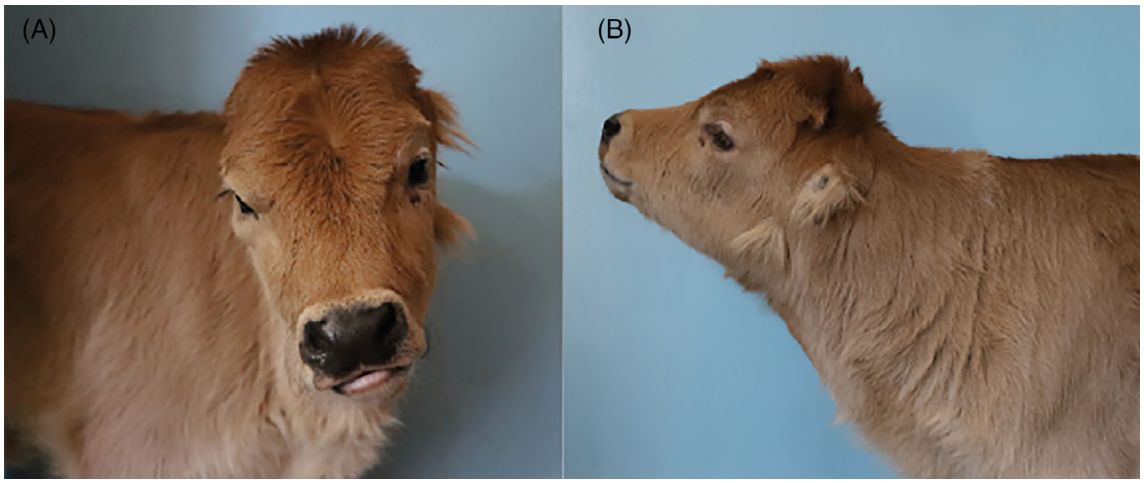


FIGURE 1 Hemifacial microsomia (HFM) in the Romagnola calf demonstrating: (A) note the abnormal conformation of the splanchnocranium with slight right deviation from the sagittal plan and (B) note that the left pinna is a rudiment of soft tissue with absence of the ear canal (aural atresia) and covered by long hair

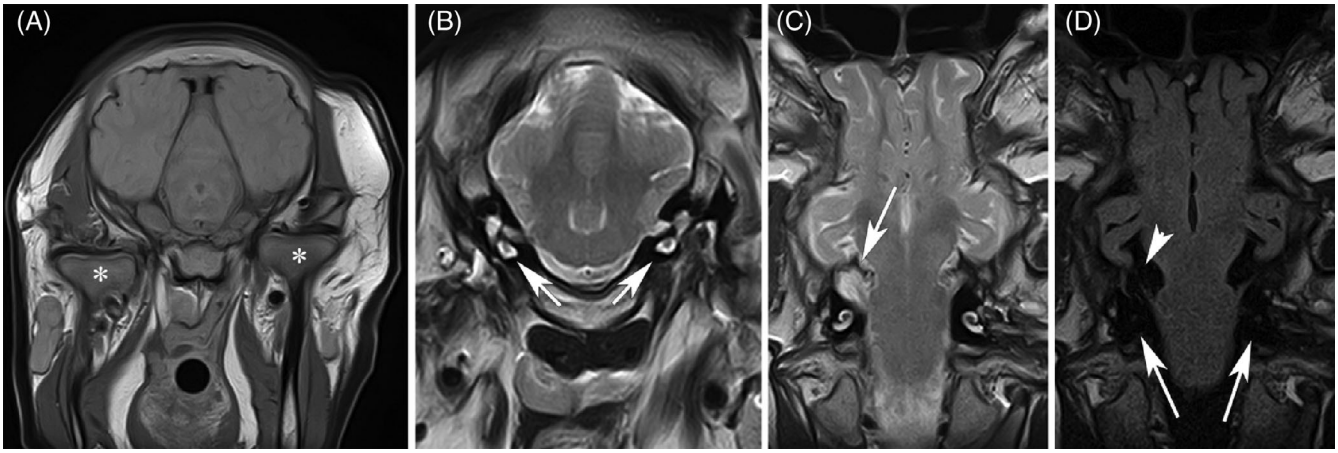


FIGURE 2 Magnetic resonance imaging of the head in the Romagnola calf with hemifacial microsomia (HFM). (A) Transverse proton density image at the level of the caudal mesencephalon. There is marked asymmetry of the temporomandibular joints (asterisks) and the surrounding soft tissues. (B) Transverse T2-weighted image at the level of the pons. There is a normal hyperintense signal of the perilymph and endolymph in both inner ears (arrows). Note the bilateral agenesia of the tympanic bullae. (C) Dorsal T2-weighted image at the level of the inner ears. An enlargement of the subarachnoidal space (meningocele) is seen on the right side at the level of the pons (arrow). (D) Dorsal fluid-attenuated inversion recovery (FLAIR) image at the same level as (C). The T2-weighted hyperintense signal from the inner ears (long arrows) and cerebrospinal fluid in the meningocele (arrow head) is suppressed

hair (Figure 1B). The neurological examination revealed reduced mental status characterized by decreased level of consciousness with listlessness and drowsiness. Notably, the calf did not respond to loud noises and hand clapping. It had a normal stance and gait. A deficit of proprioception was detected in the forelimbs.

Hematology revealed lymphocytosis (6030/mm^3; reference interval, 4250-5850/mm^3) with monocytosis (1430/mm^3) and neutrophilia (6490/mm^3; reference interval, 290-950/mm^3), and hypoproteinemia (5.68 g/dL; reference interval, 6.74-7.46 g/dL) with hypoalbuminemia (2.79 g/dL; reference interval, 3.03-3.55 g/dL). Blood samples were tested for bovine viral diarrhea virus, Schmallenberg virus, bluetongue virus, *Neospora caninum*, and *Toxoplasma gondii* using PCR and ELISA for detecting antigens and antibodies, respectively. Tests were negative for all these pathogens using both PCR and ELISA.

The calf underwent general anesthesia for magnetic resonance imaging (MRI) of the head. Magnetic resonance imaging was obtained using a 1.5 T scanner. T2-weighted images were acquired in transverse, sagittal, and dorsal planes, T1-weighted images were acquired in the transverse plane, fluid attenuated inversion recovery (T2-FLAIR) images were acquired in the dorsal plane, and proton density images were acquired in the transverse plane. Slice thickness was 3 to 4 mm, with a 10% interslice gap. Field of view was 16 to 18 cm. No contrast medium was administered. Magnetic resonance imaging revealed: asymmetry of the temporal bones and the temporomandibular joints associated with a right pontine meningocele (Figure 2A,C,D); agenesis of the right external ear canal and both tympanic bullae (Figure 2B). Moreover, on the left side, a structure resembling the innermost part of the external ear canal in shape and location was detected. There was no cavitation. T2-weighted images showed bilaterally a normally shaped, hyperintense signal of the endolymphatic and perilymphatic fluids contained in the inner ear.

Brainstem auditory evoked responses (BAERs) were examined. The signal was amplified 200 000 times, filtered with a bandwidth of 160 to 2000 Hz, and averaged 500 times. Automatic artifact rejection was used with an analysis time of 10 ms. The recording montage was vertex (noninverting input of the amplifier) and ipsilateral mastoid (inverting input). Ground electrode was inserted at the base of the neck. Recording and ground electrodes were stainless steel needles. Acoustic and bone stimuli, produced by electrical square waves of 0.1 ms with a delivery rate of 10/s, were used. Acoustic stimuli were alternating clicks of 95 dB normal hearing level (NHL) delivered monaurally using an audiometric earphone. Bone stimulation was performed with a specific transducer applied to the ipsilateral mastoid bone at a stimulus intensity of 95 dB NHL. For each ear and type of stimulation, 2 tracings were obtained and superimposed to show reproducibility of the responses. The BAERs confirmed the impaired auditory capacity with no evidence of acoustic or bone stimulation at high intensities in ear.

Three months after hospitalization the calf was euthanized because of a severe pneumonia not apparently related to the primary disease.

The calf was subsequently submitted for necropsy. Macroscopically, the left pinna was hypoplastic and the opening of the external ear canal closed by haircoat while the right pinna was absent and no anatomical remains were found. After decalcification on formalin fixed tissue, macroscopic examination was performed on cut surface having cochlea and semicircular canals aligned. On transversal cut surface, it was completely occupied by chondroid tissue. No abnormalities were detected in the inner ears and brain. Due to the absence of cerebrospinal fluid pressure after detachment of the head, the pontine meningocele observed on the right side by MRI was not detected. Additional findings were severe bronchopneumonia, complete ectopia of the spiral loop of the ascending colon, and numerous nonperforated abomasal ulcers.

Both the ear regions and brain were collected for histopathology. They were fixed in 10% neutral buffered formalin and 5 μm paraffin embedded sections were routinely stained with hematoxylin and eosin (H&E). Formalin-fixed paraffin-embedded 5 μm transverse sections of the brain were stained with H&E and Luxol-fast blue-periodic acid-Schiff methods. Histological abnormalities were not observed in both brain tissue and inner ears. The clinical and pathological findings resembled a form of HFM.

Several inbreeding loops between the unaffected parents were found in the pedigree of the calf. In light of this obvious consanguinity, the presented case of bovine HFM was hypothesized to be a rare recessively inherited variant. Therefore, WGS using the Illumina NovaSeq6000 was performed on DNA extracted from EDTA-blood of the HFM-affected calf, its dam, and from semen of its sire. The sequenced reads were mapped to the ARS-UCD1.2[18] reference genome resulting in an average read depth of approximately 18.1× in the calf, 17.9× in the dam, and 19.2× in the sire and subsequently single-nucleotide variants and small indel variants were called. The applied software and steps to process fastq-files into binary alignment map and genomic variant call format files were in accordance with the

TABLE 1 Results of variant filtering of the HFM-affected calf using the whole-genome sequence data of both parents and 4706 control genomes

Filtering step	Homozygous variants	Heterozygous variants
All variants in the affected calf	3 896 484	4 757 749
Private variants in the affected calf	104 207	1423
Private variants in the affected calf with obligatory carrier parents (protein-changing)	99 443 (245)	NA
Protein-changing private variants with obligatory carrier parents (recessive inheritance)	5	NA
Protein-changing private variants absent in both parents (de novo mutations)	NA	0

Abbreviations: HFM, hemifacial microsomia; NA, not applicable.

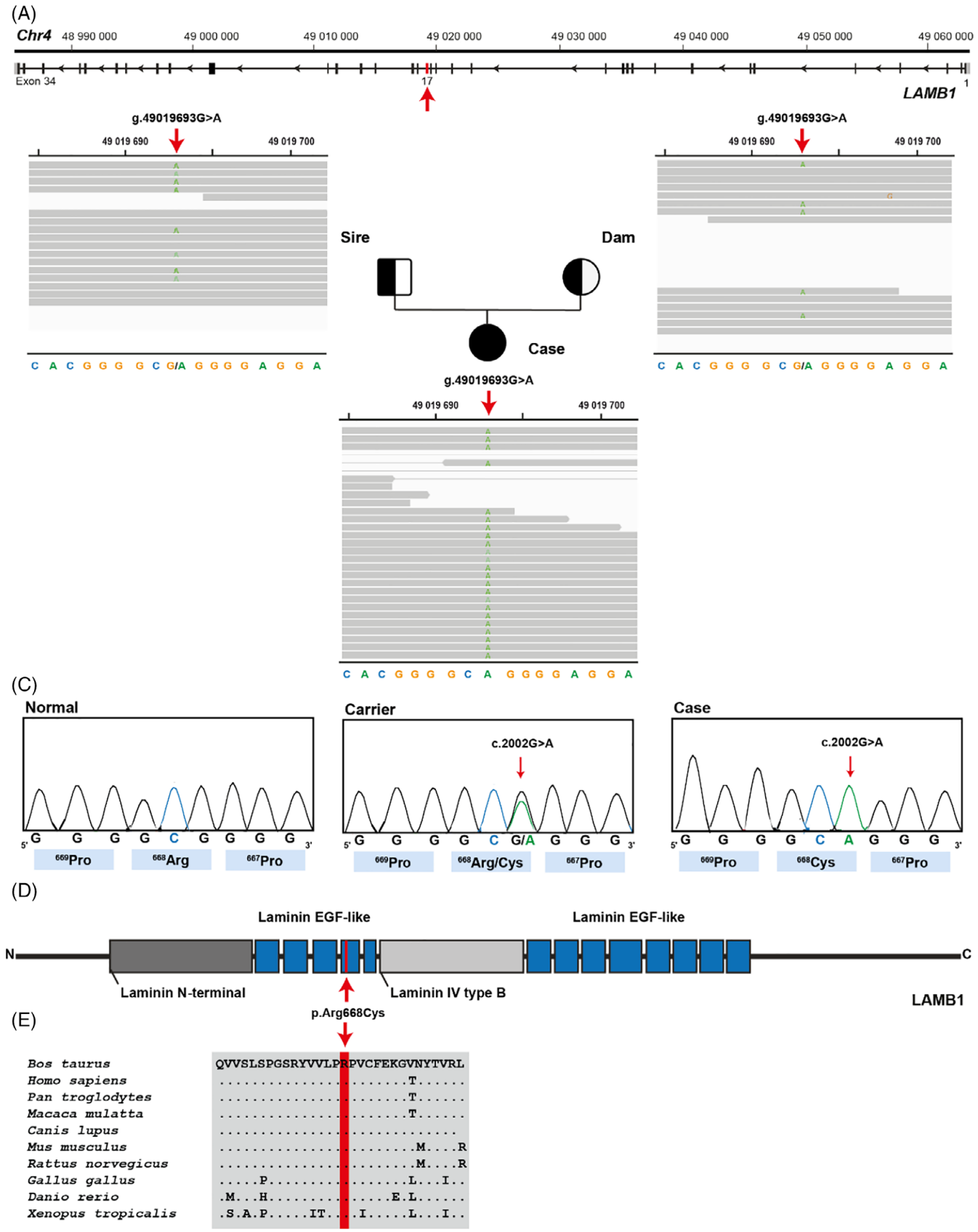


FIGURE 3 Legend on next page.

TABLE 2 Pathogenicity prediction results for the 5 homozygous protein-changing variants exclusively present in the genome of the affected calf and absent in the global control cohort of more than 4700 genomes of a variety of breeds

Gene	OMIM	Associated disorder/gene function	Protein change	Predicted effect	Provean impact and score
LAMB1	150240	Lissencephaly 5	p.Arg668Cys	Deleterious	−4.544
PDCD7	608138	Ceramide-mediated signaling	p.Pro28Leu	Neutral	−1.677
CLMN	611121	Specifically expressed at the final stage of spermatogenesis	p.Lys988Arg	Neutral	−0.310
MEX3C	611005	Phosphoproteins that bound RNA	p.Ala12Pro	Neutral	−0.087
DCC	120470	Colorectal cancer; esophageal carcinoma; Gaze palsy, familial horizontal, with progressive scoliosis, 2	p.Cys36Arg	Neutral	1.322

TABLE 3 Association of the p.Arg668Cys missense variant in *LAMB1* with the hemifacial macrosomia (HFM) phenotype in Romagnola cattle

	Genotype (R = Arg; C = Cys)		
	RR	RC	CC
HFM affected calf	0	0	1
Obligate carriers[a]	NA	2	0
Normal Romagnola control bulls	216	5	0
Normal control cattle from various breeds	4706	0	0

Abbreviation: NA, not applicable.
[a]Parents of the affected animal.

1000 Bull Genomes Project processing guidelines of run 7 (www.1000bullgenomes.com),[19] except for the trimming, which was performed using fastp.[20] Further preparation of the genomic data was done as reported earlier.[21] In order to find private variants, the genotype of the affected calf was compared with 4706 controls, including 596 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies at the Institute of Genetics of the University of Bern (Table S1) as well as 4110 genomes of a variety of breeds included in run 8 of the 1000 Bull Genomes Project.[18] The generated sequence data are publicly available in the European Nucleotide Archive (SAMEA7015114 is the sample accession number of the affected calf; SAMEA7690202 is the sample accession number of the dam and SAMEA7690203 of the sire; http://www.ebi.ac.uk/en). Integrative Genomics Viewer (IGV)[22] software version 2.0 was used for visual inspection of genome regions containing candidate variants. Assuming recessive inheritance in a trio-based approach, filtering of WGS data for homozygous coding variants present in the calf and heterozygous in the parental genomes identified 99 443 variants of which 245 were protein-changing with a predicted high or moderate impact (Table 1). These 245 variants were further investigated for their occurrence in a global control cohort of 4706 genomes of a variety of breeds, which revealed 5 remaining protein-changing variants that were exclusively homozygous in the genome of the affected calf and heterozygous in its parents (Tables 1 and S2).

Among these 5 remaining private variants, 1 single variant affects an interesting candidate gene for the observed phenotype (Figure 3A; Table 2). This homozygous variant at chr4:49019693G>A represents a missense variant in *LAMB1* (NM_001206519.1: c.2002C>T; Figure 3B,C). It alters the encoded amino acid of LAMB1 residue 668 (NP_001193448.1:p.Arg668Cys) located in the laminin epidermal growth factor (EGF)-like 4 domain (Figure 3D). Furthermore, the arginine to cysteine substitution affects an evolutionary conserved amino acid (Figure 3E) and was predicted to be deleterious[23] (Table 2). To confirm and evaluate the presence of the *LAMB1* variant, the affected genomic region was amplified by PCR and Sanger sequenced in the calf, its dam and sire. Additionally, DNA was extracted from ETDA-blood of 221 Romagnola bulls and genotyping of the *LAMB1* variant was performed. The *LAMB1* missense variant was genotyped using the following primers: 5′- GTAGATGCACGTTGTCTGCC -3′ (forward primer) and 5′- AGCCAAAACCAGACACAGACTA -3′ (reverse primer). Analyzing the sequencing data, it was confirmed that the calf was

FIGURE 3 A homozygous *LAMB1* missense variant in the HFM-affected Romagnola calf. (A) *LAMB1* gene structure showing the variant location on chromosome 4, exon 17 (red arrow). References to the bovine *LAMB1* gene correspond to the NCBI accessions NC_037331.1 (chromosome 4, ARS-UCD1.2), NM_001206519.1 (bovine *LAMB1* mRNA). (B) IGV screenshot presenting the Chr4: g. 49019693G>A variant homozygous in the affected calf (shown below) and heterozygous in both parents (top left: sire; top right: dam) revealed by whole-genome sequencing. (C) Electropherograms showing the normal, carrier, and case genotypes obtained by Sanger sequencing. (D) Schematic representation of the bovine LAMB1 protein and its functional domains obtained from the UniProt database (http://www.uniprot.org/; accession number: A0A3S5ZPX3). Laminin N-terminal domain is represented in dark gray; laminin epidermal growth factor (EGF)-like domains are represented in blue; laminin IV type B domain is represented in light gray. (E) Multiple sequence alignment of the laminin EGF -like of LAMB1 protein encompassing the region of the p.Arg668Cys variant demonstrates complete evolutionary conservation across species. Protein sequences accession numbers in NCBI for each species are NP_001193448.1 (*Bos taurus*), NP_002282.2 (*Homo sapiens*), XP_001165667.2 (*Pan troglodytes*), XP_001090393.2 (*Macaca mulatta*), XP_533089.4 (*Canis lupus*), NP_032508.2 (*Mus musculus*), XP_003750185.1 (*Rattus norvegicus*), XP_415943.3 (*Gallus gallus*), XP_002933140.2 (*Xenopus tropicalis*), NP_775382.1 (*Danio rerio*). HFM, hemifacial microsomia; IGV, Integrative Genomics Viewer

homozygous and the sire and dam heterozygous for the detected *LAMB1* variant. Furthermore, the genotyping of the 221 Romagnola bulls revealed no homozygous mutant animal and a total of 5 heterozygous carriers (1.13%; Table 3).

Variant filtering revealed no private heterozygous protein-changing variants present in the genome of the HFM-affected calf and absent in both parental genomes and in 4706 controls.

3 | DISCUSSION

In this study, a comprehensive clinical, pathologic, and genetic investigation of a deaf Romagnola calf displaying a form of congenital microtia associated with craniofacial anomalies revealed a putative genetic cause for the abnormality. In humans, 50% of microtia cases are associated with ancillary findings, mostly craniofacial anomalies.[24,25] Hemifacial microsomia is 1 of the major microtia-related diagnoses in human medicine with an incidence of 1:5600 live births[26] and is estimated as the most common birth defect of the human face, after cleft lip and cleft palate. Features of HFM include unilaterally as well as bilaterally deformity of the external ear and small ipsilateral half of the face with epibulbar dermoid and vertebral anomalies (OMIM 164210), leading to asymmetrical appearance.[27] Due to a marked phenotypic diversification, no typical clinical picture can be assigned to HFM: it might present from minor asymmetry with deformed auricle or microtia, until complete anotia, with conductive type hearing loss.[24]

A genetic origin was evaluated assuming either a recessively inherited mutation or alternatively the hypothesis of a dominant acting de novo mutation (which occurred in a single parental gamete or happened during early embryonic development of the calf) as the possible cause for this novel congenital phenotype. The trio-based WGS approach identified 5 homozygous and no heterozygous protein-changing variants exclusively present in the genome of the affected calf and absent in a global control cohort. Consequently, a de novo mutation as a possible cause for the observed phenotype seems unlikely. After in silico effect predictions just the homozygous variant affecting the fourth laminin EGF-like domain of LAMB1 was predicted to be deleterious. Moreover, within a representative control cohort of the current Italian Romagnola population, a very low allele frequency and the absence of the homozygous genotype for the deleterious allele was noticed. Considering the rarity of this coding variant, the in silico effect prediction and the known function of *LAMB1* gene, the identified variant was considered to represent the most likely genetic cause for the observed phenotype. Furthermore, it might be assumed that this pathogenic variant affecting a functional candidate gene is the most plausible explanation.

The candidate gene *LAMB1* encodes laminin subunit beta 1 belonging to laminins that are large molecular weight glycoproteins found in the basal lamina, playing an important role in cell proliferation, differentiation, migration, and adhesion.[28] They are cross-shaped heterotrimeric proteins constituted by the assembly of 3 disulfide-linked polypeptides, the α, β, and γ chains from the *LAMA*, *LAMB*, and *LAMC* families, respectively.[29] Each individual laminin subunit demonstrates a specific spatial and temporal expression pattern.[30] In mammals, there are at least 15 laminins, and LAMB1 is present in 6 of them.[31] Moreover, LAMB1 is 1 of the earliest laminin subunits expressed during embryogenesis at several sites, including neuroectoderm.[29,32] The calf in this study had several malformations deriving from neuroectoderm such as the pontine meningocele and malformations of the auricle, middle ear, and temporomandibular region. Hemifacial microsomia affects most structures of the craniofacial region that derive from the first and second brachial arches (BAs). Arising from the neuroectoderm, neural crest cells (NCCs) follow stereotypical migratory pathways and populate the BAs, and cranial NCCs (CNCCs) form the first and second BAs, which contribute to most craniofacial skeleton and connective tissues.[33] Cranial neural crest cells in the first BA form the maxilla, zygomatic bone, mandible, malleus, incus, and trigeminal nerve, which might be affected in HFM. Cranial neural crest cells of the second BA form the stapes and facial nerve.[34] Therefore the described phenotype displayed by the calf in this study, including microtia of the left ear and anotia of the right ear, absence of tympanic bullae, deafness and asymmetry of the temporal bones and the temporomandibular joints, strongly resembles human HFM. In veterinary medicine, so far, rare forms of HFM are reported only in cats.[35]

Mutations in the *LAMB1* gene have been identified and studied at a molecular level in humans (OMIM 150240) and mice (MGI 96743). Pathogenic variants affecting the human *LAMB1* are associated with autosomal recessive diseases such as: cobblestone brain malformation with congenital hydrocephalus, severe developmental delay, and an increased head circumference[36]; progressive leukoencephalopathy with seizures, ocular abnormalities, and porencephalic lesions[37]; childhood-onset epilepsy, macrocephaly, and intellectual development arrest[38]; and adult-onset leukoencephalopathy.[39] In mice, heterozygous dominant acting variants in *lamb1* are associated with dystonia-like movement disorder with brain and spinal neuronal defects,[40] while recessively inherited pathogenic variants are related with embryonic lethality between implantation and somite formation.[41] The calf presented in this study had malformation of the skull; however, no signs of leukoencephalopathy, cobblestone brain malformation, or dystonia were noticed.

This is a report of a pathogenic *LAMB1* variant in domestic animals and of the *LAMB1*-related recessively inherited form of HFM. Therefore, it represents an animal model for the understanding of similar human conditions and adds *LAMB1* to the list of candidate genes for HFM. Humans show developmental, anatomical, and physiological features of the auditory system that are more similar with cattle than with mice. Among these, the fact that compared to mice, cattle and humans can hear at birth. The cattle model thus fits better than the mouse model for understanding human hearing diseases.

In conclusion, this study provides a DNA-based diagnostic test that enables selection against the identified pathogenic variant in Romagnola cattle.

ACKNOWLEDGMENT

No funding was received for this study. The authors acknowledge the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure. The WGS data are available under

the study accession no. PRJEB18113 at the European Nucleotide Archive (www.ebi.ac.uk/ena; sample accessions SAMEA7015114 [affected calf], SAMEA7690202 [dam], and SAMEA7690203 [sire]). We thank Nathalie Besuchet-Schmutz for expert technical assistance.

CONFLICT OF INTEREST DECLARATION

Authors declare no conflicts of interest.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION

Authors declare no IACUC or other approval was needed.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Joana G. P. Jacinto https://orcid.org/0000-0002-6438-7975
Marco Bernardini https://orcid.org/0000-0002-8572-0271
Maria Teresa Mandara https://orcid.org/0000-0003-1501-6163
Marilena Bolcato https://orcid.org/0000-0002-0605-3344
Arcangelo Gentile https://orcid.org/0000-0002-6091-8978
Cord Drögemüller https://orcid.org/0000-0001-9773-522X

REFERENCES

1. Suutarla S, Rautio J, Ritvanen A, Ala-Mello S, Jero J, Klockars T. Microtia in Finland: comparison of characteristics in different populations. *Int J Pediatr Otorhinolaryngol*. 2007;71(8):1211-1217.
2. Luquetti DV, Heike CL, Hing AV, Cunningham ML, Cox TC. Microtia: epidemiology and genetics. *Am J Med Genet A*. 2012;158A(1):124-139.
3. Mastroiacovo P, Corchia C, Botto LD, Lanni R, Zampino G, Fusco D. Epidemiology and genetics of microtia-anotia: a registry based study on over one million births. *J Med Genet*. 1995;32(6):453-457.
4. Qiao R, He Y, Pan B, et al. Understanding the molecular mechanisms of human microtia via a pig model of HOXA1 syndrome. *Dis Model Mech*. 2015;8(6):611-622.
5. Tischfield MA, Bosley TM, Salih MAM, et al. Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat Genet*. 2005;37(10):1035-1037.
6. Alasti F, Sadeghi A, Sanati MH, et al. A mutation in HOXA2 is responsible for autosomal-recessive microtia in an Iranian family. *Am J Hum Genet*. 2008;82(4):982-991.
7. Brown KK, Viana LM, Helwig CC, et al. HOXA2 haploinsufficiency in dominant bilateral microtia and hearing loss. *Hum Mutat*. 2013;34(10):1347-1351.
8. Bicknell LS, Bongers EMHF, Leitch A, et al. Mutations in the pre-replication complex cause Meier-Gorlin syndrome. *Nat Genet*. 2011;43(4):356-360.
9. Guernsey DL, Matsuoka M, Jiang H, et al. Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nat Genet*. 2011;43(4):360-365.
10. Fenwick AL, Kliszczak M, Cooper F, et al. Mutations in CDC45, encoding an essential component of the pre-initiation complex, cause Meier-Gorlin syndrome and craniosynostosis. *Am J Hum Genet*. 2016;99(1):125-138.
11. Vetro A, Savasta S, Russo Raucci A, et al. MCM5: a new actor in the link between DNA replication and Meier-Gorlin syndrome. *Eur J Hum Genet*. 2017;25(5):646-650.
12. Edwards SJ, Gladwin AJ, Dixon MJ. The mutational spectrum in Treacher Collins syndrome reveals a predominance of mutations that create a premature-termination codon. *Am J Hum Genet*. 1997;60(3):515-524.
13. Schaefer E, Collet C, Genevieve D, et al. Autosomal recessive POLR1D mutation with decrease of TCOF1 mRNA is responsible for Treacher Collins syndrome. *Genet Med*. 2014;16(9):720-724.
14. Sanchez E, Laplace-Builhé B, Mau-Them FT, et al. POLR1B and neural crest cell anomalies in Treacher Collins syndrome type 4. *Genet Med*. 2020;22(3):547-556.
15. Tekin M, Öztürkmen Akay H, Fitoz S, et al. Homozygous FGF3 mutations result in congenital deafness with inner ear agenesis, microtia, and microdontia. *Clin Genet*. 2008;73(6):554-565.
16. He S, Zhang Z, Sun Y, et al. Genome-wide association study shows that microtia in Altay sheep is caused by a 76 bp duplication of HMX1. *Anim Genet*. 2020;51(1):132-136.
17. Koch CT, Bruggmann R, Tetens J, Drögemüller C. A non-coding genomic duplication at the HMX1 locus is associated with crop ears in highland cattle. *PLoS One*. 2013;8(10):e77841.
18. Rosen BD, Bickhart DM, Schnabel RD, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020;9(3):1-9.
19. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. *Annu Rev Anim Biosci*. 2019;7:89-102.
20. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890.
21. Häfliger IM, Wiedemar N, Švara T, et al. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim Genet*. 2020;51(3):382-390.
22. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the integrative genomics viewer. *Cancer Res*. 2017;77(21):e31-e34.
23. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015;31(16):2745-2747.
24. Paul MA, Opyrchał J, Knakiewicz M, Jaremków P, Bajtek J, Chrapusta A. Hemifacial microsomia review: recent advancements in understanding the disease. *J Craniofac Surg*. 2020;31(8):2123-2127.
25. Hartzell LD, Chinnadurai S. Microtia and related facial anomalies. *Clin Perinatol*. 2018;45(4):679-697.
26. Wolford LM, Perez DE. Surgical management of congenital deformities with temporomandibular joint malformation. *Oral Maxillofac Surg Clin North Am*. 2015;27(1):137-154.
27. Gougoutas AJ, Singh DJ, Low DW, Bartlett SP. Hemifacial microsomia: clinical features and pictographic representations of the OMENS classification system. *Plast Reconstr Surg*. 2007;120(7):112e-113e.
28. Lee J, Gross JM. Laminin beta1 and gamma1 containing laminins are essential for basement membrane integrity in the zebrafish eye. *Invest Ophthalmol Vis Sci*. 2007;48(6):2483-2490.
29. Roediger M, Miosge N, Gersdorff N. Tissue distribution of the laminin β1 and β2 chain during embryonic and fetal human development. *J Mol Histol*. 2010;41(2-3):177-184.
30. Sharif KA, Li C, Gudas LJ. cis-acting DNA regulatory elements, including the retinoic acid response element, are required for tissue specific laminin B1 promoter/lacZ expression in transgenic mice. *Mech Dev*. 2001;103(1-2):13-25.
31. Sharif KA, Baker H, Gudas LJ. Differential regulation of laminin b1 transgene expression in the neonatal and adult mouse brain. *Neuroscience*. 2004;126(4):967-978.

32. Baldarelli RM, Smith CM, Finger JH, et al. The mouse Gene Expression Database (GXD): 2021 update. *Nucleic Acids Res*. 2021;49(D1):D924-D931.
33. Minoux M, Rijli FM. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development*. 2010;137(16):2605-2621.
34. Johnson JM, Moonis G, Green GE, Carmody R, Burbank HN. Syndromes of the first and second branchial arches, part 1: embryology and characteristic defects. *AJNR Am J Neuroradiol*. 2011;32(1):14-19.
35. Song RB, Kent M, Glass EN, Davis GJ, Castro FA, de Lahunta A. Hemifacial microsomia in a cat. *Anat Histol Embryol*. 2017;46(5):497-501.
36. Radmanesh F, Caglayan AO, Silhavy JL, et al. Mutations in LAMB1 cause cobblestone brain malformation without muscular or ocular abnormalities. *Am J Hum Genet*. 2013;92(3):468-474.
37. Tonduti D, Dorboz I, Renaldo F, et al. Cystic leukoencephalopathy with cortical dysplasia related to LAMB1 mutations. *Neurology*. 2015;84(21):2195-2197.
38. Okazaki T, Saito Y, Hayashida T, et al. Bilateral cerebellar cysts and cerebral white matter lesions with cortical dysgenesis: expanding the phenotype of LAMB1 gene mutations. *Clin Genet*. 2018;94:391-392.
39. Yasuda R, Yoshida T, Mizuta I, et al. Adult-onset leukoencephalopathy with homozygous LAMB1 missense mutation. *Neurol Genet*. 2020;6(4):4-7.
40. Liu YB, Tewari A, Salameh J, et al. A dystonia-like movement disorder with brain and spinal neuronal defects is caused by mutation of the mouse laminin β1 subunit, Lamb1. *Elife*. 2015;4:e11102.
41. Miner JH, Li C, Mudd JL, Go G, Sutherland AE. Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. *Development*. 2004;131(10):2247-2256.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Jacinto JGP, Häfliger IM, Bernardini M, et al. A homozygous missense variant in laminin subunit beta 1 as candidate causal mutation of hemifacial microsomia in Romagnola cattle. *J Vet Intern Med*. 2021;1-8. doi:10.1111/jvim.16316

***De novo* stop-lost germline mutation in *FGFR3* causes severe chondrodysplasia in the progeny of a Holstein bull**

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, methodology, writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.12934
John Wiley and Sons License Number:	5236050301928

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12934

De novo stop-lost germline mutation in *FGFR3* causes severe chondrodysplasia in the progeny of a Holstein bull

I. M. Häfliger*, A. Letko*, L. Murgiano*,†,‡ and C. Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland. †Unit of Animal Genomics, GIGA-R and Faculty of Veterinary Medicine, University of Liège, Liège 4000, Belgium. ‡Department of Clinical Sciences and Advanced Medicine, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Summary

Fifteen cases of chondrodysplasia characterized by disproportionate dwarfism occurred in the progeny of a single Holstein bull. A *de novo* mutation event in the germline of the sire was suspected as cause. Whole-genome sequencing revealed a single protein-changing variant in the stop codon of *FGFR3* gene on chromosome 6. Sanger sequencing of EDTA blood proved that this variant occurred *de novo* and segregates perfectly with the observed phenotype in the affected cattle family. *FGFR3* is an important regulator gene in bone formation owing to its key role in the bone elongation induced by FGFR3-dimers. The detected paternally inherited stop-lost variant in *FGFR3* is predicted to add 93 additional amino acids to the protein's C-terminus. This study provides a second example of a dominant *FGFR3* stop-lost variant as a pathogenic mutation of a severe form of chondrodysplasia. Even though *FGFR3* is known to be associated with dwarfism and growth disorders in human and sheep, this study is the first to describe *FGFR3*-associated chondrodysplasia in cattle.

Keywords calf survival, dwarfism, genetic disorder, monogenic

Chondrodysplasia (CD) represents a developmental bone defect occurring owing to disturbed endochondral osteogenesis disorders of the development of bone and cartilage. In livestock, CD is characterized by disproportionate dwarfism and manifests with short and stumpy limbs, also reported as achondroplasia (Boegheim *et al.* 2017). Congenital forms of proportionate and disproportionate dwarfism associated with different genes and different modes of inheritance have been described in a number of cattle breeds showing varying viable or lethal phenotypes (OMIA 000004-9913, OMIA 000187-9913, OMIA 000311-9913, OMIA 001485-9913, OMIA 001686-9913, OMIA 001926-9913, OMIA 001985-9913, OMIA 002171-9913). For example, the Bulldog calf syndrome, a lethal form of bovine CD characterized by a generalized shortening of long bones, is associated either with six known dominantly inherited protein-changing variants in the *COL2A1* gene (Häfliger *et al.* 2019) or with two different semidominant loss-of-function variants in the *ACAN* gene (Cavanagh *et al.* 2007). Interestingly, individuals heterozygous for these *ACAN* variants exhibit disproportionate dwarfism (Harper *et al.* 1998; Cavanagh *et al.* 2007). In disproportionate dwarfism, there is a height reduction in one or several body parts and body proportions are altered (Boegheim *et al.* 2017). Viable recessive forms of disproportionate dwarfism are associated with variants in the *EVC2* gene described in Japanese Brown cattle (Takeda *et al.* 2002) and Tyrolean Grey cattle (Murgiano *et al.* 2014), whereas a similar disorder occurring in American Angus cattle is caused by a nonsense variant introducing a premature stop codon in the *PRKG2* gene (Koltes *et al.* 2009).

A single sire was exclusively used on a Holstein purebred breeding farm in Germany to naturally mate almost all cows and heifers (Drögemüller *et al.* 2006). Out of 60 pregnancies delivered full-term, 15 CD-affected calves of both sexes were identified at birth. The observed disease phenotype described before (Drögemüller *et al.* 2006) includes a mild shortening of the upper jaw (brachygnathia superior) and an abnormal stature with movement disabilities owing to severe skeletal shortening of the limbs and hyperextension of the joints (Fig. 1; Videos S1 and S2). Furthermore, an examination of the profound shortened long bones of CD-affected calves revealed a rarefication of primary spongiosa. Interestingly, there was no indication for irregularly arranged chondrocytes of epiphyseal plates (Drögemüller *et al.* 2006), as reported in other forms of

Address for correspondence

C. Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland.
E-mail: cord.droegemueller@vetsuisse.unibe.ch

Accepted for publication 12 March 2020

bovine dwarfism or chondrodysplasia. Whereas a germline mutation was hypothesized in Drögemüller *et al*. (2006), we present herein our successful efforts to unravel the causative variant.

A total of 34 samples of genomic DNA extracted from EDTA-stabilized blood were available including eight CD-affected (four female, four male) and 13 apparently normal calves, 12 out of 21 dams and the sire of all offspring. We obtained WGSs from four family members: two CD-affected offspring, their common sire and the dam of one of these two affected calves. Therefore, either Illumina TruSeq Nano (case 1) or TruSeq PCR-free libraries (case 2, sire, dam) was prepared for sequencing paired-end reads on either an Illumina HiSeq 2000 (case 1) or a NovaSeq 6000 instrument (Illumina, San Diego, CA, USA). The resulting fastq-files were mapped to the cow reference genome assembly ARS-UCD1.2 and single nucleotide and small indel variants were called according to the guidelines of the 1000 Bull Genomes consortium (Hayes & Daetwyler 2019; http://www.1000bullgenomes.com/). NCBI annotation release 106 was used to predict their functional effects as described previously (Häfliger *et al*. 2019). Initially, the genome of a single CD-affected calf (case 1) was sequenced at approximately 15× coverage but we failed to detect a plausible candidate variant causing the disorder. Therefore, we opted for trio-based sequencing of another CD-affected calf at approximately 15× coverage and both of its parents at approximately 20× coverage. The WGS data of the four animals are stored in the European Nucleotide Archive under project accession no. PRJEB18113 [sample accessions SAMEA19846918 (case 1), SAMEA5415502 (case 2), SAMEA5415503 (dam) and SAMEA5415501 (sire)].

Genome-wide filtering for sequence variants that were present only in the genome of case 2 and absent in the genomes of its parents, as well as in 395 controls of various breeds that had been sequenced in the course of other ongoing studies (Table S1), resulted in 47 variants passing the quality filter and representing putative *de novo* sequence variants. Only a single heterozygous variant was predicted as protein-changing in *FGFR3* (Chr 6: 116 767 863C>A), a gene which is associated with multiple types of skeletal dysplasia including achondroplasia in man (OMIM 134934). The variant affecting the second base of the stop codon of bovine *FGFR3* (NM_174318.3: c.2408G>T) was predicted to extend the sequence at the C-terminal end with 93 additional amino acids [XM_024992994.1: p.(Ter803Leuext*93); Appendix S1]. Integrative Genomics Viewer (Robinson *et al*. 2011) was used for visual inspection of the identified variant in exon 17 of the bovine *FGFR3* gene and confirmed that the variant was absent in both parents without any evidence for low-level mosaicism in one of the parents (Fig. 2a). Visual re-evaluation of the previously sequenced genome of the initially sequenced CD-affected calf confirmed that the variant was also present in a heterozygous state (Fig. 2a). Most likely the slightly lower read depth of approximately 6× for that specific genome region in case 1 explains the original overlooking of this variant. This *de novo* heterozygous mutation in *FGFR3* was considered a plausible candidate causal variant. Further support was given by comparing the 47 variants with the 1000 Bull Genomes run 7 (Hayes & Daetwyler 2019) as the variant in our study that eliminated the *FGFR3* stop codon was the only variant not found in international variant catalog including 3103 control genomes. In addition, to rule out a possible recessively inherited disease-causing variant, we filtered the WGS data for this alternative scenario. However, this revealed no indication for the presence of a homozygous protein-changing variant shared by the two affected cases.

Genotyping of all 34 available family members by Sanger sequencing of a PCR product flanking the variant (forward primer 5′-GCATGACCTGTGAGTGCG-3′, reverse primer 5′-GACCCGGTGATCATGGC-3′) revealed that all eight CD-affected calves were heterozygous for the stop-lost variant (Fig. 2b), whereas the remaining 26 animals were all homozygous wt. We suggest that the observed alleles in the CD-affected offspring were due to a *de novo* mutation event, as all genotyped parents were homozygous for the wt allele. A manual haplotype segregation analysis with 12 informative single-nucleotide variants flanking the *FGFR3* variant confirmed the assumed paternal origin of the mutation (Appendix S2).

In sum, the *de novo* variant in *FGFR3* is perfectly associated with dominantly inherited chondrodysplasia in this family of Holstein cattle. Neither next-generation nor Sanger sequencing showed any presence of the identified variant in the sire, who clearly carried only the wt alleles in DNA originating from blood. Considering that approximately 25% of offspring were affected, we assume that the sire is a germline mosaic for the variant but unfortunately no semen was available to confirm.

Figure 1 Two Holstein calves with chondrodysplasia used for whole-genome sequencing. Note the severe skeletal shortening of the limbs and hyperextension of the joints.

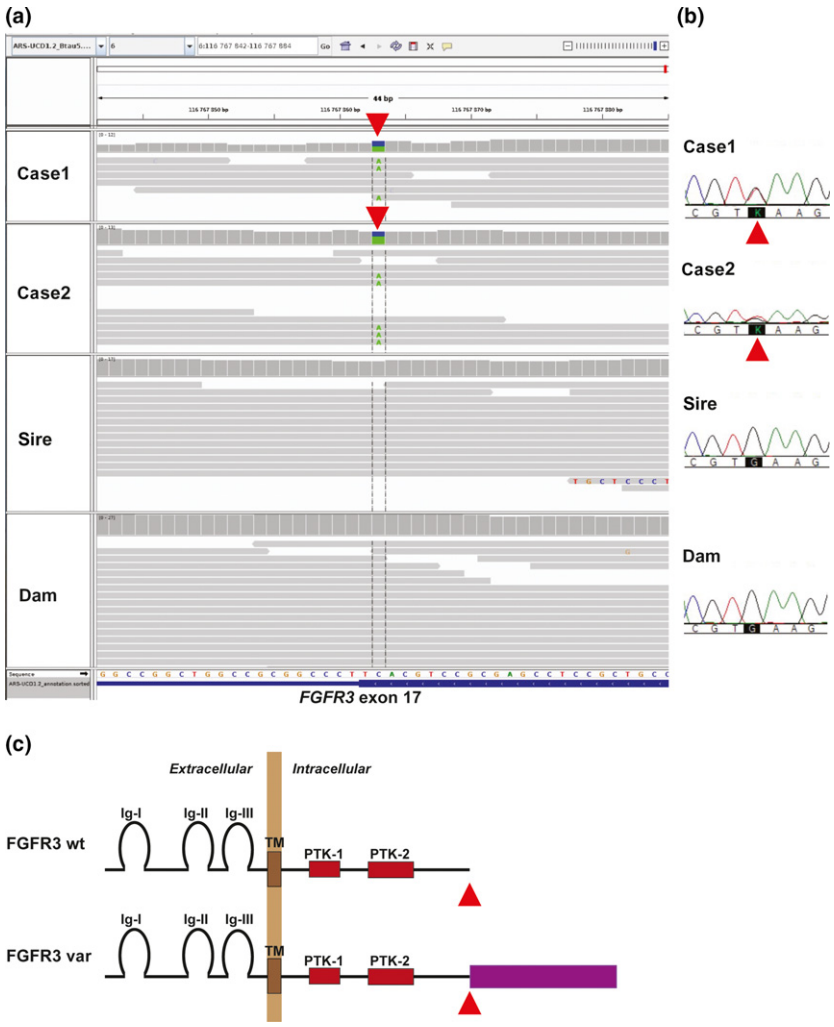


Figure 2 Features of the *de novo FGFR3* stop-lost variant. The variant site is indicated by the red triangles. (a) Screenshot of Integrative Genomics Viewer indicating a heterozygous variant in the stop codon of the *FGFR3* gene in the two cases, whereas the sire and dam of case 2 do not show the variant. (b) Sanger sequencing of the variant. Note that the electropherograms show that the mutant T allele is present in heterozygous form in the chondrodysplasia-affected offspring only. (c) Schematic representation of the FGFR3 protein structure, according to Laederich & Horton (2012). Note the significant elongation of the mutant protein (FGFR3 var).

Fibroblast growth factor receptor 3 (FGFR3) is a member of the tyrosine kinase receptor family, which plays an important role in the signaling pathway of the fibroblast growth factors (FGF). FGF receptors, such as FGFR3, contain an extracellular domain with either two or three immunoglobulin (Ig)-like domains, a transmembrane domain and a cytoplasmic tyrosine kinase domain (Fig. 2c). Normally, FGFR3-dimers are built by FGFs binding to FGFR3 and activating the dimerization of FGFR3, which results in restrained (or limited) bone elongation (Moosa & Wollnik 2016). As FGFs/FGFRs are key players in both endochondral and intramembranous bone development, naturally occurring mutations cause severe developmental disorders (Moosa & Wollnik 2016). Variants in human *FGFR3* are associated with various forms of achondroplasia and dwarfism representing different phenotypes, ranging from shortened limbs to lethal thanatophoric dysplasia (OMIM 134934; Horton *et al.* 2007; Moosa & Wollnik 2016). Most reported *FGFR3* variants represent dominantly inherited gain-of-function mutations, which lead to an increased FGFR3 dimerization and subsequent abnormal regulation of growth plate chondrocytes typically seen in chondrodysplasia (Horton *et al.* 2007; Moosa & Wollnik 2016). In domestic animals, so far only a single missense variant in ovine *FGFR3* has been reported, causing a semidominantly inherited skeletal overgrowth phenotype known as spider lamb syndrome in Suffolk sheep (OMIA 001703-9940; Beever *et al.* 2006). The variant presented herein is the first *FGFR3* mutation reported in cattle. Interestingly, similar to our variant, Rousseau *et al.* (1995) found three different heterozygous base substitutions in the stop codon of *FGFR3* in human patients with thanatophoric dysplasia type I (OMIM 187600). This severe short-limb dwarfism syndrome that is usually lethal in the perinatal period is caused by a *de novo* stop-lost mutation which is expected to give rise to a protein elongated by 141 amino acids, resulting in a highly hydrophobic domain with an alpha-helix structure at the C-terminal end of the full-length protein. The same characteristics were found in our study when analyzing the additional 93 amino acids of the variant bovine FGRF3 protein *in silico* using PSIPRED

Protein Analysis Workbench (Buchan & Jones 2019; Appendix S1). The elongated proteins are supposed to block the dimerization of FGFR3 proteins and disturb its receptor function. Therefore, the identified stop-lost variant represents the most likely pathogenic variant owing to a gain in FGFR3 function of the elongated protein. Functional validation experiments would be needed to confirm the biological impact of such kind of variants on the resultant transcripts or proteins. For example, the detected *FGFR3* variant of high impact could have been validated through direct sequencing of transcripts and mass spectrometry to detect elongated and misfolded FGFR3 proteins. Unfortunately, no suitable samples were available in this study, and therefore, such experiments were not possible.

In conclusion, the disproportional dwarfism observed in the progeny of a single Holstein bull represents an *FGFR3*-related form of chondrodysplasia. This highlights the potential risk of unlimited use of natural-service sires in a single breeding herd because of the potential of a protein-changing germline mutation. The study provides a second example of a dominant *FGFR3* stop-lost variant as a pathogenic mutation of a severe form of chondrodysplasia in a mammalian species. This case could expand knowledge about the phenotypic spectrum of *FGFR3* variants regarding especially the negative effects of an elongated FGRF3 protein.

Acknowledgements

The Next Generation Sequencing Platform of the University of Bern is acknowledged for performing the whole-genome re-sequencing experiments and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computing infrastructure. The authors are grateful to Holger Lönnecker and Nathalie Besuchet-Schmutz for technical assistance. This study was partially funded by the Swiss National Science Foundation.

References

Beever J.E., Smit M.A., Meyers S.N., Hadfield T.S., Bottema C., Albretsen J. & Cockett N.E. (2006) A single-base change in the tyrosine kinase II domain of ovine FGFR3 causes hereditary chondrodysplasia in sheep. *Animal Genetics* **37**, 66–71.

Boegheim I.J., Leegwater P.A., van Lith H.A. & Back W. (2017) Current insights into the molecular genetic basis of dwarfism in livestock. *Veterinary Journal* **224**, 64–75.

Buchan D.W.A. & Jones D.T. (2019) The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Resarch* **47**, W402–7.

Cavanagh J.A.L., Tammen I., Windsor P.A., Bateman J.F., Savarirayan R., Nicholas F.W. & Raadsma H.W. (2007) Bulldog dwarfism in Dexter cattle is caused by mutations in ACAN. *Mammalian Genome* **18**, 808–14.

Drögemüller C., Starke A., Schmidbauer S. & Wohlsein P. (2006) Kongenitaler disproportionaler Zwergwuchs bei Deutschen Holsteins. *Tierärztliche Praxis Ausgabe Grosstiere Nutztiere* **34**, 148–54.

Häfliger I.M., Behn H., Freick M., Jagannathan V. & Drögemüller C. (2019) A *COL2A1 de novo* variant in a Holstein bulldog calf. *Animal Genetics* **50**, 113–4.

Harper P.A.W., Latter M.R., Nicholas F.W., Cook R.W. & Gill P.A. (1998) Chondrodysplasia in Australian Dexter cattle. *Australian Veterinary Journal* **76**, 199–202.

Hayes B.J. & Daetwyler H.D. (2019) 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual Review of Animal Biosciences* **7**, 89–102.

Horton W.A., Hall J.G. & Hecht J.T. (2007) Achondroplasia. *Lancet* **370**, 162–72.

Koltes J.E., Mishra B.P., Kumar D. *et al.* (2009) A nonsense mutation in cGMP-dependent type II protein kinase (PRKG2) causes dwarfism in American Angus cattle. *Proceedings of the National Academy of Sciences USA* **106**, 19250–5.

Laederich M.B. & Horton W.A. (2012) FGFR3 targeting strategies for achondroplasia. *Expert Reviews in Molecular Medicine* **14**, e11.

Moosa S. & Wollnik B. (2016) Altered FGF signalling in congenital craniofacial and skeletal disorders. *Seminars in Cell & Developmental Biology* **53**, 115–25.

Murgiano L., Jagannathan V., Benazzi C., Bolcato M., Brunetti B., Muscatello L.V., Dittmer K., Piffer C., Gentile A. & Drögemüller C. (2014) Deletion in the *EVC2* gene causes chondrodysplastic dwarfism in Tyrolean Grey cattle. *PLoS ONE* **9**, e94861.

Robinson J.T., Thorvaldsdottir H., Winckler W., Guttman M., Lander E.S., Getz G. & Mesirov J.P. (2011) Integrative genomics viewer. *Nature Biotechnology* **29**, 24–6.

Rousseau F., Saugier P., Lemerrer M., Munnich A., Delezoide A.L., Maroteaux P., Bonaventure J., Narcy F. & Sanak M. (1995) Stop codon *Fgfr3* mutations in thanatophoric dwarfism type-1. *Nature Genetics* **10**, 11–2.

Takeda H., Takami M., Oguni T. *et al.* (2002) Positional cloning of the gene LIMBIN responsible for bovine chondrodysplastic dwarfism. *Proceedings of the National Academy of Sciences USA* **99**, 10549–54.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Accession numbers of the 395 publicly available genome sequences used as control genomes.

Appendix S1. FGFR3 sequences of the wt and variant (var) cDNA (NM_174318.3) and protein (XM_024992994.1).

Appendix S2. Haplotype analysis of the *FGFR3* genome region on chromosome 6.

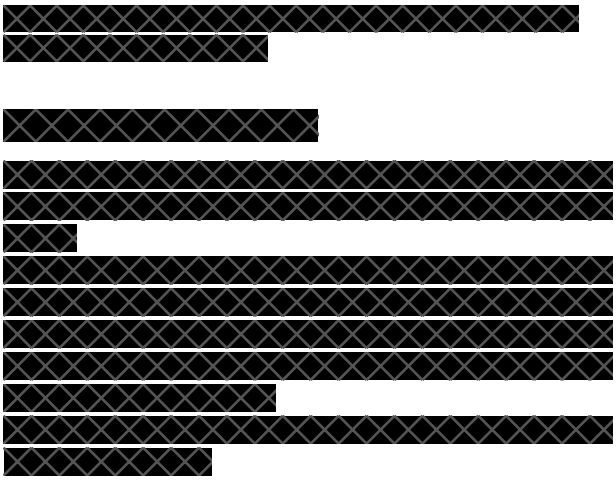
Video S1. Movement restrictions of a chondrodysplasia (CD)-affected calf (case 1). Note the hyperextension of the joints, resulting in inability to walk.

Video S2. Gait abnormalities of another CD-affected calf. Note discontinuous walking direction.

A *COL2A1 de novo* variant in a Holstein bulldog calf

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, co-writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.12735
John Wiley and Sons License Number:	5236051032621

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics



doi: 10.1111/age.12735

A *COL2A1 de novo* variant in a Holstein bulldog calf

Irene Monika Häfliger*, Holger Behn†, Markus Freick‡§, Vidhya Jagannathan* and Cord Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland; †Landesuntersuchungsanstalt für das Gesundheits- und Veterinärwesen Sachsen, 04158 Leipzig, Germany; ‡Veterinary Practice Zettlitz, 09306 Zettlitz, Germany; §Faculty for Agriculture/Environment/Chemistry, HTW Dresden – University of Applied Sciences, 01326 Dresden, Germany

Accepted for publication 13 September 2018

Background: Bulldog (BD) calf syndrome is a lethal form of chondrodysplasia characterized by a generalized shortening of long bones, also known as bovine achondrogenesis type II (OMIA 001926-9913).[1] Over the past 18 years, BD has been reported in 1% of purebred calves from a French Holstein sire, in 4% of crossbred calves from a French Charolais bull, in 12% of purebred calves from a Danish Holstein sire, in 21% of purebred calves from a German Holstein sire and in an isolated case born from German Holstein parents.[1–5] Today, a total of five independent pathogenic dominant-acting *COL2A1* variants explaining these last-mentioned BD cases, including a splice site variant and four missense variants, have been identified.[2–5] The *COL2A1* gene encodes the alpha-1 chain of type II collagen, and in humans, non-synonymous *COL2A1* variants that disrupt the Gly-x-y structural motif essential for the assembly of the collagen triple-helix cause hypochondrogenesis/achondrogenesis type II (ACG2). This human dominant-inherited condition shares strong similarities with the phenotypic features observed in bovine BD calf syndrome, and the disease-causing *COL2A1* variants are either inherited from mosaic parents, mainly sires, or arises *de novo* in the developing embryo (MIM 200610).

Own analysis: A female stillborn calf showing severe malformation born after normal gestation of 265 days was noticed from a mating of a fifth-calf Holstein cow and a progeny-tested Holstein bull used intensively for artificial insemination (AI). Based on the records of both the cow's owner and the AI station, no further evidence for other affected offspring of either parents could be revealed. Macroscopic necropsy results revealed a morphology similar to Holstein calves with disproportional lethal chondrodysplasia as described before (Fig. S1). A variant in the *COL2A1* gene was therefore a possible cause. The calf tested negative for BVDV, Schmallenberg virus and *Coxiella burnetii* using RT-PCR, and bacterial cultures for *E. coli*, *Brucella* and *Salmonella* spp. were negative.

The genome of the affected calf was sequenced at ~13× coverage using 2 × 150 bp reads on a HiSeq 3000 instrument. The sequencing data were mapped to the UMD3.1 reference genome and submitted to the European Nucleotide Archive under project accession no. PRJEB18113 (sample accession SAMEA4644751). A genome-wide a total of 7 890 388 single nucleotide variants (SNVs), small insertion/deletion variants (indels), were called with respect to the reference assembly as described.[6] Due to the lethal effect, we hypothesized that most likely a protein-changing variant would be responsible for the observed disease phenotype. The sequenced calf carried 31 046 protein-changing variants, which included not only missense SNVs but also nonsense SNVs, indels in protein-coding regions and splice site variants (Table S1), including 1665 variants that are not contained in the known *Bos taurus* dbSNP 51 dataset. A single heterozygous private variant Chr5:32 476 808G>A in the *COL2A1* gene is predicted to cause substitution of a glycine residue with serine (NM_001001135.3:c.3166G>A; NP_001001135.2:p.Gly1056Ser). Glycine at this position is conserved across vertebrates (Fig. S1).

Sanger sequencing of PCR products using genomic DNA samples extracted from the semen of the sire and the blood of the dam revealed no evidence that the parents of the affected calf were obviously germline or somatic mosaic, as the identified variant in *COL2A1* was absent in both parents (Fig. S1).

Comments: The identified *COL2A1* missense variant, like most of the alterations responsible for human ACG2, disrupts the invariant GXY structural motif (with any amino acid at the second and third positions) necessary for perfect triple-helix formation and could thus lead to extensive overmodification, intracellular retention and reduced secretion of type II collagen, as previously established in the human form of the disorder.[7] These predominantly glycine to non-serine residue substitutions exclusively create more severe phenotypes, such as achondrogenesis type II and hypochondrogenesis, and correspond to the bovine *COL2A1* allele most likely causing BD calf syndrome in the presented case.

Furthermore, it could be concluded that the identified *COL2A1* p.Gly1056Ser missense variant occurred *de novo.* It is highly likely that the mutation occurred during early embryonic development of the affected calf. This is exactly the case for one (p.Gly720Ser) of the four BD calf syndromes causing missense variants that were previously reported.[4] For Holstein cattle it is important to recognize that, even within a single breed, phenotypically indistinguishable cases of BD calf syndrome show notable allelic heterogeneity.

Acknowledgements: We thank the Masterrind GmbH for providing semen of the sire.

References

1 Agerholm J.S., Arnbjerg J. & Andersen O. (2004) *J Vet Diagn Invest* **16**, 293–8.
2 Daetwyler H.D. *et al.* (2014) *Nat Genet* **46**, 858–65.
3 Agerholm J.S. *et al.* (2016) *BMC Vet Res* **12**, 100.
4 Bourneuf E. *et al.* (2017) *Sci Rep* **7**, 11466.
5 Reinartz S. *et al.* (2017) *BMC Genomics* **18**, 762.
6 Murgiano L. *et al.* (2016) *BMC Vet Res* **12**, 276.
7 Vissing H.J. *et al.* (1989) *Biol Chem* **264**, 18265–7.

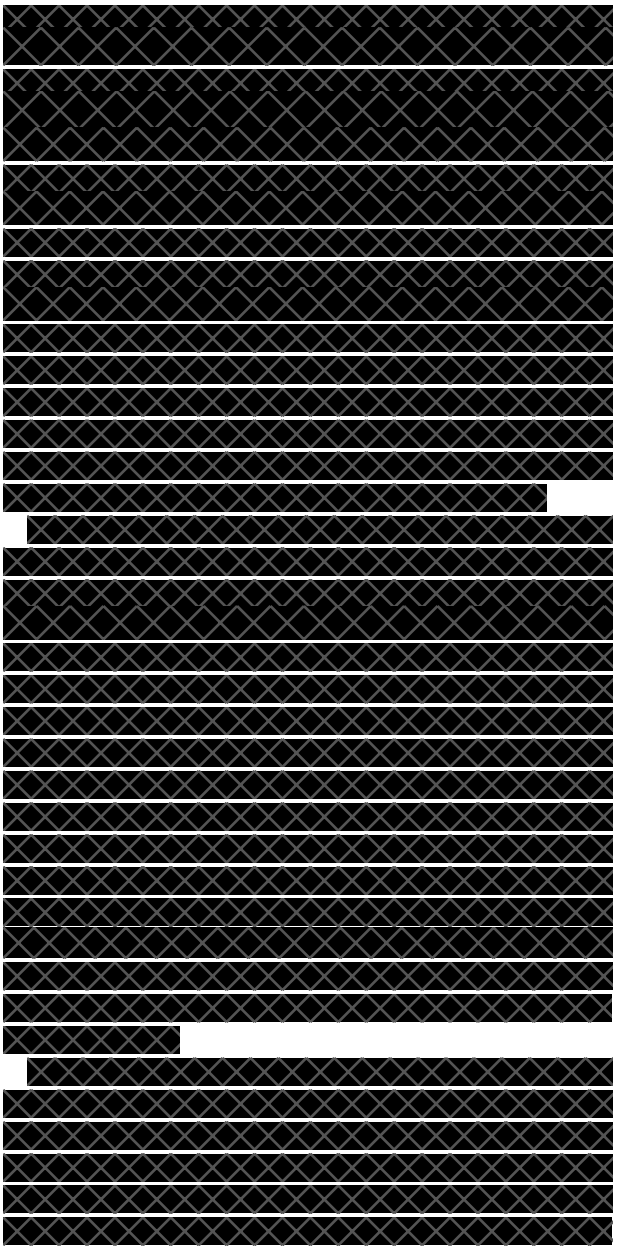
Correspondence: C. Drögemüller
(cord.droegemueller@vetsuisse.unibe.ch)

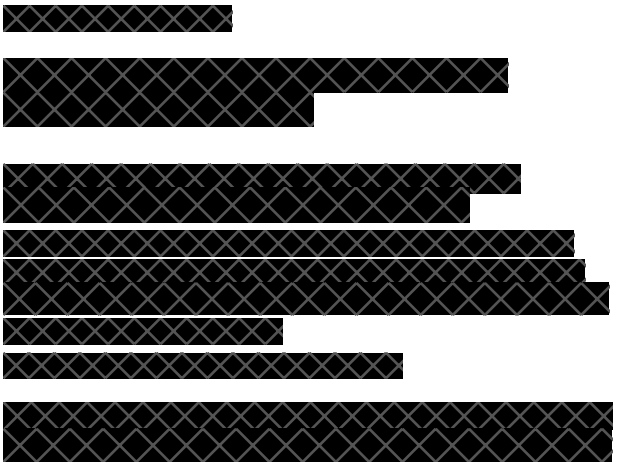
Supporting information

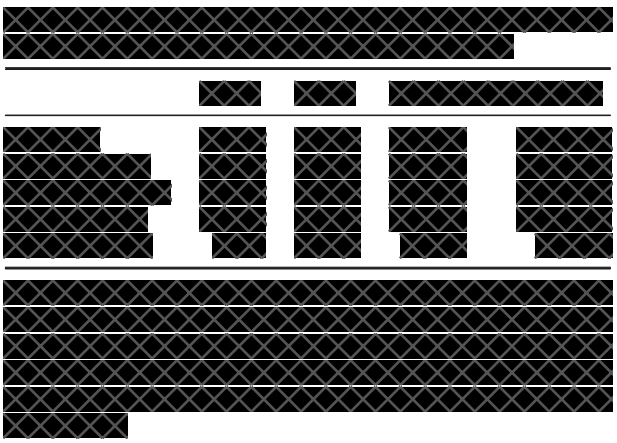
Additional supporting information may be found online in the Supporting Information section at the end of the article:

Figure S1 Identification and characterization of the *COL2A1*:g.32476808G>A variant in a Holstein calf with Bulldog calf syndrome.

Table S1 Protein-changing variants of the sequenced animal.







A 6.7 kb deletion in the *COL2A1* gene in a Holstein calf with achondrogenesis type II and perosomus elumbis

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.1111/age.13033
John Wiley and Sons License Number:	5236050702081

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

BRIEF NOTES

doi: 10.1111/age.13033

A 6.7 kb deletion in the *COL2A1* gene in a Holstein calf with achondrogenesis type II and perosomus elumbis

Joana G. P. Jacinto*,† , Irene M. Häfliger† , Arcangelo Gentile* , Cord Drögemüller† and Marilena Bolcato*

*Department of Veterinary Medical Sciences, University of Bologna, Ozzano Emilia, Bologna 40064, Italy; †Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland

Accepted for publication 02 December 2020

Background

Bovine achondrogenesis type II is also known as bulldog calf (BD) and is caused by a congenital chondrodysplasia characterized by disproportionate growth of bones, resulting in a shortened and compressed body, mostly because of the reduced length of the spine and the long bones of the limbs.[1] Moreover, severe facial dimorphisms, e.g. palathoschisis and shortening of the viscerocranium, are present.[1] Recessively inherited variants in the *ACAN* gene are associated with the so-called lethal Dexter BD type[2] (OMIA 001271-9913), whereas dominant inherited *COL2A1* variants are related to forms of bovine achondrogenesis type II or BD (OMIA 001926-9913). The latter may be inherited from mosaic parents[3–6] or due post-zygotic *de novo* muations.[5,7,8] So far, a total of seven pathogenic variants in the *COL2A1* have been reported.[3–8]

Own analysis

A stillborn purebred Holstein male calf was delivered after dystocia. Gross pathology findings revealed a phenotype resembling the bovine chondrodysplasia type II with the additional presence of perosomus elumbis (Fig. S1). WGS was performed using genomic DNA obtained from the ear tissue of the calf as described before.[7] The sequenced reads were mapped to the ARS-UCD1.2 cattle reference genome,[9] resulting in an average read depth of approximately 19.8×. The WGS data of the case can be found in the European Nucleotide Archive under the sample accession no. SAMEA7690227. A deleterious variant in the *COL2A1* gene was hypothesized to be causal. Therefore, INTEGRATIVE GENOMICS VIEWER software[10] was used for visual inspection of variants in the region of the *COL2A1* gene on chromosome 5. A single pair of reads indicated the presence of a 6.7 kb-sized deletion, and the sequence coverage within the deleted segment is apparently depleted compared with the flanking regions (Fig. S1). The deletion spanning 19 coding exons of the *COL2A1* gene was subsequently evaluated by performing a multiplex PCR across both breakpoints, revealing an additional PCR product only in the affected calf (Fig. S1). The two obtained PCR products across the breakpoints represent the wt allele, and thereby we confirmed that the case showing a third PCR product was indeed heterozygous for the suspected structural variant. Sanger sequencing revealed the precise breakpoints of the heterozygous deletion from position 32 301 911 located in intron 25 to 32 308 589 located within exon 45. The 6679 bp deletion includes the entire sequence of 18 exons (26–44) plus the first 36 nucleotides of exon 45 (Fig. S1).

Comments

The heterozygous deletion is predicted to affect a large portion of the *COL2A1* gene. Two possible scenarios could cause the observed phenotype: either haploinsufficiency or co-expression of a significantly truncated protein. The *COL2A1* gene has a probability of loss-of-function intolerance score of 1, meaning that it clearly falls into the class of loss-of-function haploinsufficient genes.[11] Consequently, the observed phenotype could be explained by the non-expression of the mutant allele. This has been speculated recently as a cause for a BD case showing a heterozygous 3.5 kb deletion encompassing 10 exons of *COL2A1*.[8] On the other hand, the variant detected herein is deleting 1413 bp of the coding sequence representing an in-frame deletion that affects 438 residues of the triple helical region of the COL2A1 protein. Owing to the lack of suitable material, it remains unclear whether this shortened transcript encoding a significantly truncated protein is expressed or not. Nonetheless, the invariant Gly-x-y structural motif is mandatory for perfect triple-helix formation and could thus lead to extensive overmodification. It could be speculated that the variant allele disrupts the triple-helical region of alpha 1 (II) chain causing a dominant-negative effect similar to most of the alterations responsible for achondrogenesis/hypochondrogenesis type II in human patients (OMIM 200610).

Interestingly, so far phenotypically indistinguishable cases of BD calf syndrome in Holstein cattle have shown notable allelic heterogeneity. This report presents another large structural variant affecting the *COL2A1* gene causing a novel form of bovine achondrogenesis type II that occurred in combination with the perosomus elumbis (PE) phenotype (OMIA 000789-9913). Bovine PE has been known as a congenital entity for a long time and shows a certain morphological variation among cases in Holstein cattle, but so far no molecular cause has been reported.[12] Based on our findings, we postulate the COL2A1 gene as a possible candidate gene for PE in cattle.

Acknowledgements: We thank Nathalie Besuchet Schmutz for expert technical assistance and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure. This study was partially funded by the Swiss National Science Foundation.

References

1 Agerholm J.S. *et al.* (2004) *J Vet Diagn Invest* **16**, 293–8.
2 Cavanagh J.A. *et al.* (2007) *Mamm Genome* **18**, 808–14.
3 Daetwyler H.D. *et al.* (2014) *Nat Genet* **46**, 858–65.
4 Agerholm J.S. *et al.* (2016) *BMC Vet Res* **12**, 100.
5 Bourneuf E. *et al.* (2017) *Sci Rep* **7**, 11466.
6 Reinartz S. *et al.* (2017) *BMC Genomics* **18**, 762.
7 Häfliger I.M. *et al.* (2019) *Anim Genet* **50**, 113–4.
8 Jacinto J.G.P. *et al.* (2020) *Acta Vet Scand* **62**, 49.
9 Rosen B.D. *et al.* (2020) *Gigascience* **9**, 1–9.
10 Robinson J.T. *et al.* (2017) *Cancer Res* **77**, e31–4.
11 Karczewski K.J. *et al.* (2020) *Nature* **581**, 434–43.
12 Agerholm J.S. *et al.* (2014) *BMC Vet Res* **10**, 227.

Correspondence: C. Drögemüller
(cord.droegemueller@vetsuisse.unibe.ch)

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.
Figure S1. *COL2A1* deletion in a Holstein calf with achondrogenesis type II and persomus elumbis.

A large deletion in the *COL2A1* gene expands the spectrum of pathogenic variants causing bulldog calf syndrome in cattle

Journal:	Acta Veterinaria Scaninavica
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.1186/s13028-020-00548-w

Acta Veterinaria Scandinavica

CASE REPORT Open Access

A large deletion in the *COL2A1* gene expands the spectrum of pathogenic variants causing bulldog calf syndrome in cattle

Joana Gonçalves Pontes Jacinto[1,2], Irene Monika Häfliger[2], Anna Letko[2], Cord Drögemüller[2*†] and Jørgen Steen Agerholm[3†]

Abstract

Background: Congenital bovine chondrodysplasia, also known as bulldog calf syndrome, is characterized by disproportionate growth of bones resulting in a shortened and compressed body, mainly due to reduced length of the spine and the long bones of the limbs. In addition, severe facial dysmorphisms including palatoschisis and shortening of the viscerocranium are present. Abnormalities in the gene *collagen type II alpha 1 chain* (*COL2A1*) have been associated with some cases of the bulldog calf syndrome. Until now, six pathogenic single-nucleotide variants have been found in *COL2A1*. Here we present a novel variant in *COL2A1* of a Holstein calf and provide an overview of the phenotypic and allelic heterogeneity of the *COL2A1*-related bulldog calf syndrome in cattle.

Case presentation: The calf was aborted at gestation day 264 and showed generalized disproportionate dwarfism, with a shortened compressed body and limbs, and dysplasia of the viscerocranium; a phenotype resembling bulldog calf syndrome due to an abnormality in *COL2A1*. Whole-genome sequence (WGS) data was obtained and revealed a heterozygous 3513 base pair deletion encompassing 10 of the 54 coding exons of *COL2A1*. Polymerase chain reaction analysis and Sanger sequencing confirmed the breakpoints of the deletion and its absence in the genomes of both parents.

Conclusions: The pathological and genetic findings were consistent with a case of "bulldog calf syndrome". The identified variant causing the syndrome was the result of a de novo mutation event that either occurred post-zygotically in the developing embryo or was inherited because of low-level mosaicism in one of the parents. The identified loss-of-function variant is pathogenic due to *COL2A1* haploinsufficiency and represents the first structural variant causing bulldog calf syndrome in cattle. Furthermore, this case report highlights the utility of WGS-based precise diagnostics for understanding congenital disorders in cattle and the need for continued surveillance for genetic disorders in cattle.

Keywords: Chondrodysplasia, Congenital, Malformation, Precision medicine, Rare disease, Type II collagenopathy, Whole-genome sequencing

Background

The bulldog calf syndrome (BDS) is a congenital form of bovine chondrodysplasia affecting bones with endochondral osteogenesis. In its most severe form, this syndrome is lethal [1]. The BDS is often exemplified by the Dexter BDS type [2], which is linked to abnormalities in the *aggrecan (ACAN)* gene [3]. However other BDS types,

*Correspondence: cord.droegemueller@vetsuisse.unibe.ch
†Cord Drögemüller and Jørgen Steen Agerholm contributed equally to this work
[2] Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstr. 109a, 3001 Bern, Switzerland
Full list of author information is available at the end of the article

which share gross morphology features with Dexter type, are associated with abnormalities in other genes and occur in different cattle breeds. Abnormalities in the *collagen type II alpha 1 chain* (*COL2A1*) gene causing BDS have been reported several times during the last 15 years (achondrogenesis/hypochondrogenesis type II in *Bos taurus*; OMIA 001926-9913; https://omia.org/OMIA001926/9913/). The purpose of this study was to report a variant in the *COL2A1* gene leading to BDS and provide an overview of the phenotypic and allelic heterogeneity of *COL2A1*-related BDS.

Case presentation

A stillborn Holstein male calf with a body weight of 18.1 kg was aborted at gestation day 264 (normal gestation 281 days (mean)). The pregnancy was the result of insemination with semen of a purebred Holstein sire on a Holstein dam. The parents were not related within at least four generations. The calf had moderate autolysis and was frozen at − 20 °C before submission for necropsy and was examined after thawing.

The calf had generalized disproportionate dwarfism resembling a case of BDS (Fig. 1). The body appeared shortened and compact. The limbs showed bilateral symmetric shortening, which especially affected the bones proximal to the phalanges, giving the limbs a compact appearance. The phalanges were slightly rotated medially. The limbs were sawed longitudinally, which confirmed the irregular development of diaphysis and the presence of enlarged chondroid epiphyses without ossification centers (Fig. 2a). Radiological examination prior to sawing revealed normally structured phalangeal bones, but otherwise bones were only seen as irregular diaphyseal segments that could only be identified based on their location (Fig. 2b). Vertebrae had a similar appearance with enlarged chondroid epiphyses and irregular diaphyses. The head had dysplasia of the viscerocranium with shortening of the maxillary bones, palatoschisis, protrusion of the tongue and doming of the calvarium (Fig. 3). Longitudinal sawing of the head through the midline revealed that the direction of the brain axis was elevated due to the abnormally shaped neurocranium (Fig. 3a). Radiological examination highlighted the abnormally shaped bones (Fig. 3b). The thorax was narrow and of reduced volume and was mostly occupied by an enlarged malformed heart. The heart malformation consisted of bilateral ventricular dilation, muscular hypertrophy of the right ventricular wall and dilation of the pulmonary trunk. The lung was hypoplastic and nonaerated. The abdomen was dilated with eventration of intestinal segments and the liver appeared indurated. Due to the level of autolysis and freezing artefacts, histopathology was not performed.

Genetic analysis

Whole-genome sequencing using the NovaSeq 6000 (illumina) was performed at a read depth of ~ 26× using DNA extracted from skin and cartilage from the ear of

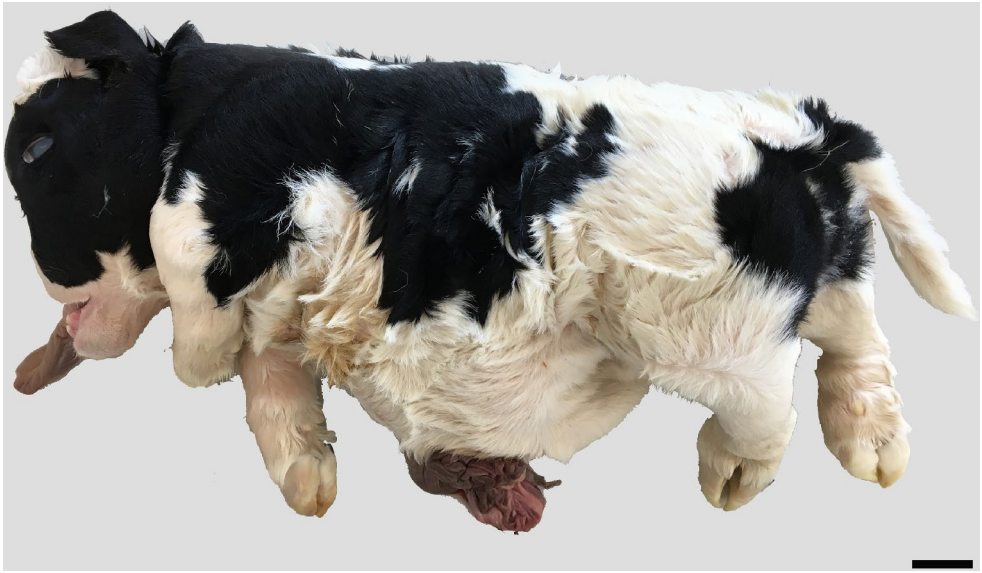


Fig. 1 Calf affected by the OMIA 001926-9913 bulldog calf syndrome. Note the short compact body and limbs and the malformed head. Segments of the intestine protruded through a defect in the ventral abdominal midline, probably as a result of insufficient space within the abdomen due to the short spine. Bar: 5 cm

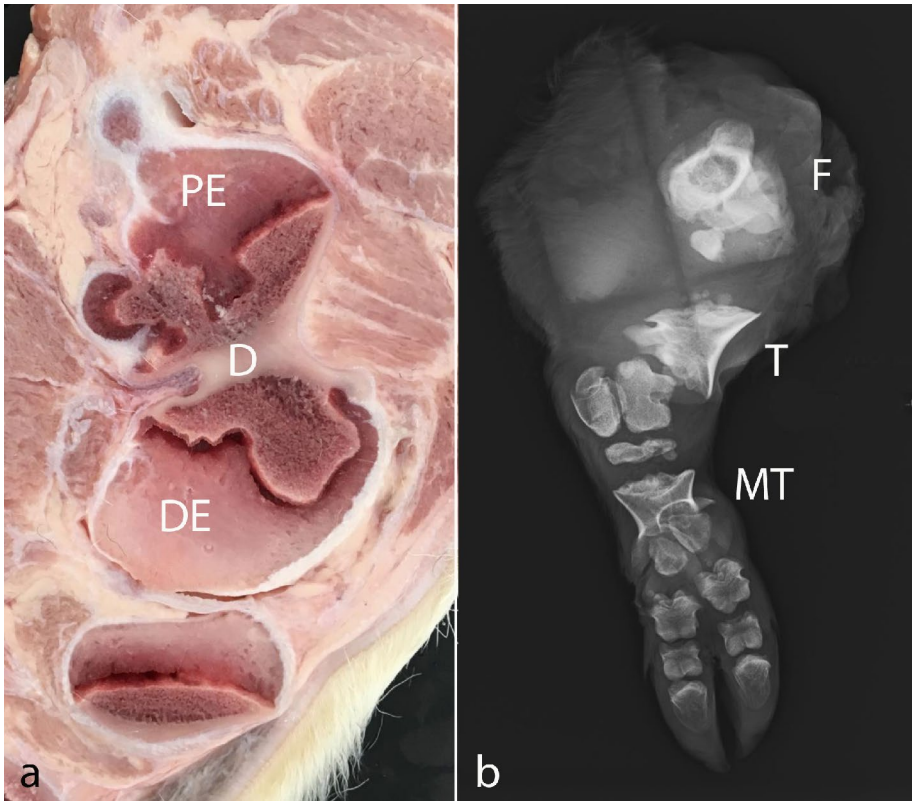


Fig. 2 Lesions in the bones of the limbs. **a** Parasagittal longitudinal section through the femur showing a dysplastic irregular shaped diaphysis (D) bordered by enlarged chondroid proximal and distal epiphyses (PE and DE, respectively). **b** Radiographic image of a hind limb illustrating the severe bone abnormalities of most bones. Of the femur (F), tibia (T) and metatarsal bones (MT), only their irregularly shaped diaphysis are seen and these bones were only identifiable by their location. The phalangeal bones were well developed and of normal shape

the calf. The generated sequences were mapped to the ARS-UCD1.2 reference genome, and single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format (GVCF) files were in accordance with the latest 1000 Bull Genomes Project processing guidelines (www.1000bullgenomes.com) [4]. Furthermore, CombineGVCFs and CatVariants of GATK v3.8 [5] were used to combine the GVCF files and the VariantFiltration tool of GATK was used to give the variants quality labels based on the standard GATK best practices. Lastly, functional impacts were annotated using SNPEFF v4.3 [6] by integrating the information from the NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/). With the resulting GVCF, including all individual variants and their functional predictions, filtering for private variants was performed. We compared the genotypes of the calf with 494 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies. The WGS data of the case can be found on ENA under the sample accession number SAMEA6528902, while a comprehensive list with all ENA accession numbers is shown in Additional file 1. A total of 20 private protein-changing single-nucleotide or short indel variants with a moderate or high predicted impact, located within 19 different genes or loci, were identified (Additional file 2). This list included no variants in *COL2A1*, the most likely candidate gene for the observed BDS phenotype. Therefore, Integrative Genomics Viewer (IGV) [7] software was used for visual inspection of the genome region containing *COL2A1* on chromosome 5. A heterozygous 3513 base pair (bp) sized deletion from position 32,303,127 to 32,306,640 spanning 10 coding exons of the *COL2A1* gene leading to haploinsufficiency of the encoded collagen type II alpha 1 chain protein was observed (Fig. 4). The heterozygous gross deletion variant in *COL2A1* is predicted to lead to a loss-of-function of the encoded collagen type II alpha 1 chain protein and was not observed in any of the 494 cattle genomes used for comparison. Therefore, this variant was further investigated as a potentially causative variant for the observed phenotype.

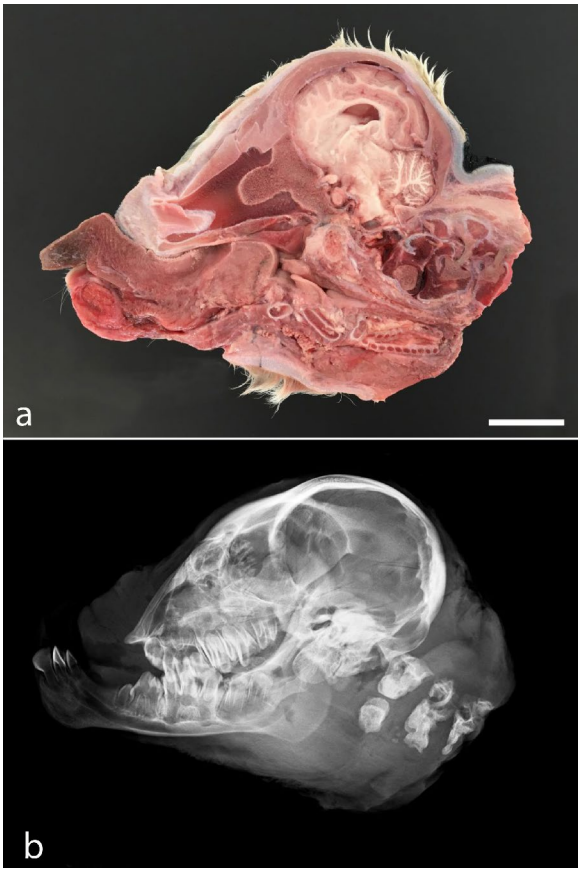


Fig. 3 Gross morphology and radiographic findings in the head. **a** Midline section showing the shortened dysplastic face and the dorso-caudal angling of the brain stem (indicated by black lines). The section through the spine is slightly parasagttital, so the spinal cord is not seen. Frozen specimen. Bar: 5 cm. **b** Same specimen as in **a** highlightening the abnormally shaped bones

To evaluate whether the deletion in *COL2A1* occurred de novo, the affected genomic region was amplified by polymerase chain reaction (PCR) and Sanger sequenced using DNA of the calf and both parents. Genomic DNA was extracted from EDTA blood and semen of the dam and sire respectively, and compared to the calf's DNA. PCR products were obtained by using primers flanking the detected *COL2A1* deletion (forward: 5′-CAGGGG ATGGGTCTTCCT-3′ and reverse: 5′-GCGTTAGAG AGGGAGACAGG-3′) and subsequently sequenced on an ABI3730 capillary sequencer (Thermofisher, Darmstadt, Germany). Only with the DNA of the affected calf, a PCR product of 128 bp could be amplified, whereas for both parents the amplification failed. Sanger sequencing of the obtained amplicon confirmed the previously identified breakpoints in combination with the insertion of a 10-bp segment fused in-between (chr5:g.32303127_32306640delinsTCTGGGGAGC).

Discussion and conclusions

Based on the morphology of the presented BDS case, a causative genetic variation in the *COL2A1* gene was suspected. As for humans, the morphology of BDS in cattle vary widely both in the overall gross morphology and in bone morphology as exemplified in Fig. 5. It appears that cases of BDS due to abnormalities in the *COL2A1* gene share a common morphology that separates them from at least some other types of bovine BDS, although few types of bovine BDS have been characterised to the molecular level. BDS cases due to abnormalities in the *COL2A1* gene are delivered at term or during the last 3 weeks of gestation. The affected calves have a significantly reduced body weight with a mean of 22.3 kg (variation 16.3–27.5 kg) for 11 Holstein cases [1, 8, 9]. The body and limbs are short and compressed with the digits being almost half of normal size, but normally shaped. The long bones of the limbs and the vertebrae have small irregular diaphyses and enlarged chondroid epiphyses. The viscerocranium is dysplastic with palatoschisis, the neurocranium doomed causing dorso-caudal rotation of the brain, the heart is malformed due to the narrow-spaced thorax, the lungs compressed and the liver with signs of chronic stasis. Cases of BDS that share this morphology, should be suspected of having a defect in the *COL2A1* gene; a suspicion that is helpful when analysing WGS data. However, in this case, filtering for private variants in *COL2A1* did not lead to the detection of a private single-nucleotide or short indel variant. Consequently, the genome data was visually inspected for the presence of structural variants in the gene that allowed the detection of a heterozygous gross deletion. It was assumed that it had occurred either post-zygotically in the developing embryo or was inherited from a parent having low-level mosaicism. The former seems to be more likely as amplification of the mutant allele failed in the examined tissues of both parents, especially because the germline of the sire was analysed by extracting DNA from semen. This means that the *COL2A1* deletion observed in heterozygous state in the affected offspring was most likely absent in the genome of both parents. Therefore, we can assume that the identified mutation arose indeed de novo in the developing embryo explaining this isolated case.

This pathogenic variant is predicted to affect a large portion of the *COL2A1* gene leading to haploinsufficiency. Recent large data from human genome sequencing studies presented in the Genome Aggregation Database (gnomAD) [10] showed that the probability of loss-of-function intolerance score for *COL2A1* was 1 meaning that *COL2A1* falls into the class of loss-of-function haploinsufficient genes. Collagens are normally extracellular structural proteins involved in formation of connective tissue structure. The highly conserved

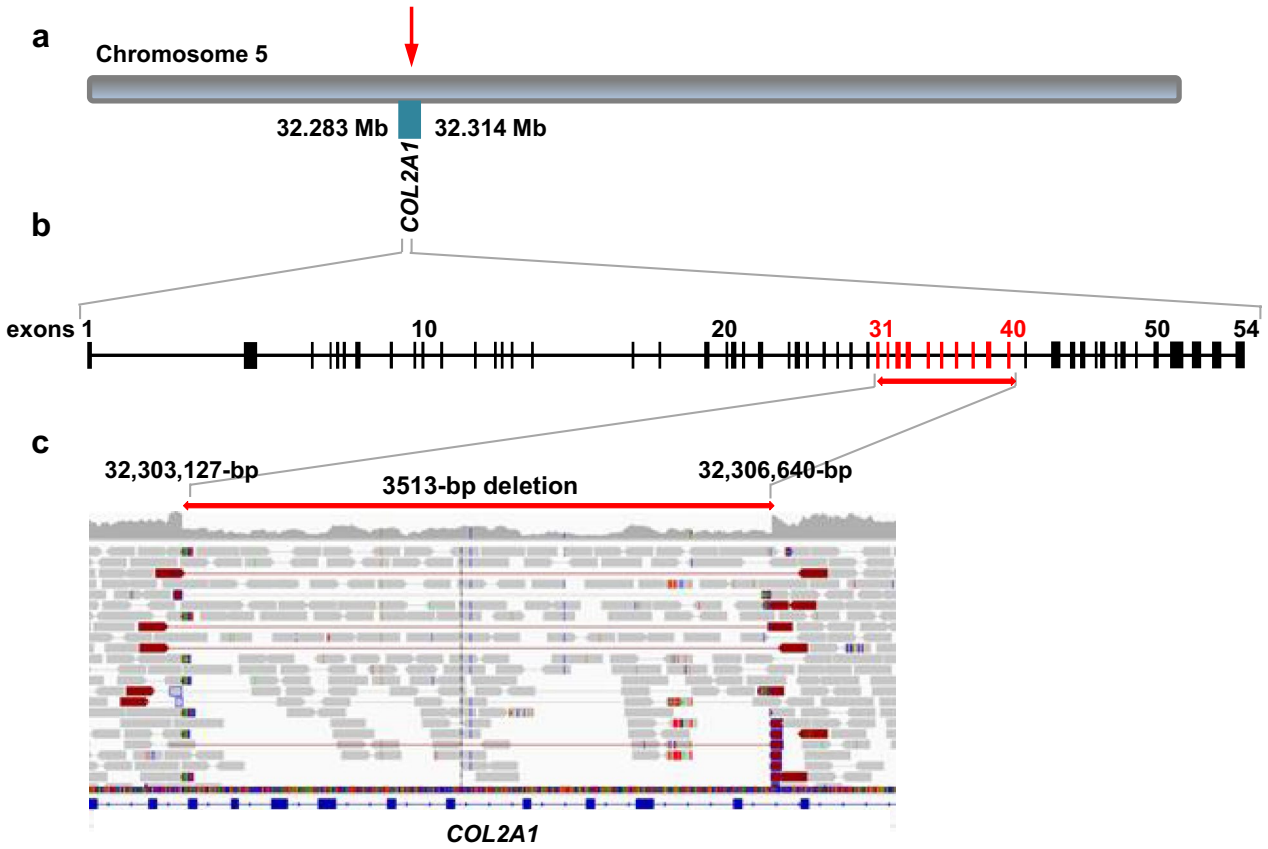


Fig. 4 Schematic diagram of the *COL2A1* gene showing the pathogenic variant causing BDS in the affected calf. **a** Location of the bovine *COL2A1* gene and causal variant indicated by the arrow on chromosome 5 of the ARS-UCD1.2 bovine genome assembly. **b** Schematic illustration of the bovine *COL2A1* gene with its 54 annotated exons. **c** Integrative Genomics Viewer screenshot in the region of the heterozygous 3513-bp deletion in combination affecting exons 31 to 40 of the *COL2A1* gene. Note the drop in coverage and the truncated read-alignments at the deletion breakpoints. Sequence reads coloured in dark red indicate the deletion as read-pairs that map farther apart on the reference genome than the average insert size of the sequencing library

sequence predominantly consists of repeated three amino acids with glycine (Gly) followed by two other amino acids (Gly-x-y, where x and y can be any amino acid) but glycine being mandatory for the tight packing of the polyproline II type helices within the triple helix [11] (Fig. 6). We assume that the pathogenic variant reported in this study disrupted the triple-helical region of alpha 1 (II) chain and caused a dominant-negative effect similar to most of the alterations responsible for achondrogenesis/hypochondrogenesis type II (OMIM 200610; https://www.omim.org/entry/200610) in human patients. In man, variants in *COL2A1* are associated with 15 different phenotypes exclusively following dominant inheritance (OMIM 120140).

Interestingly, the OMIA 001926-9913 BDS type occurs either as de novo or inherited from a mosaic parent [12]. Mosaic sires have been found to transmit the dominant genetic abnormality to their offspring at rates ranging from 1 to 21% [12, 13] reflecting at what fetal developmental stage the gene change occurred. As it cannot be predicted if the abnormality is occurring de novo or if it is transmitted from a parent, cases must be analyzed in detail to prevent the birth of large numbers of defective offspring, in particular if the abnormality is transmitted from breeding sire with high generic merit used for artificial breeding.

A total of six independent pathogenic dominant variants in *COL2A1*, considered to be responsible for BDS, have been previously identified [8, 9, 12–14] (Table 1). All these variants involve a single nucleotide; five out of the six reported variants represent missense variants that cause a change in a glycine residue disrupting the Gly-x–y structural motif essential for the assembly of the collagen triple-helix.

This is the first report of a large deletion in the *COL2A1* gene associated with BDS. The previous reported single nucleotide variants were missense and splicing. The relevance of this case report is to show

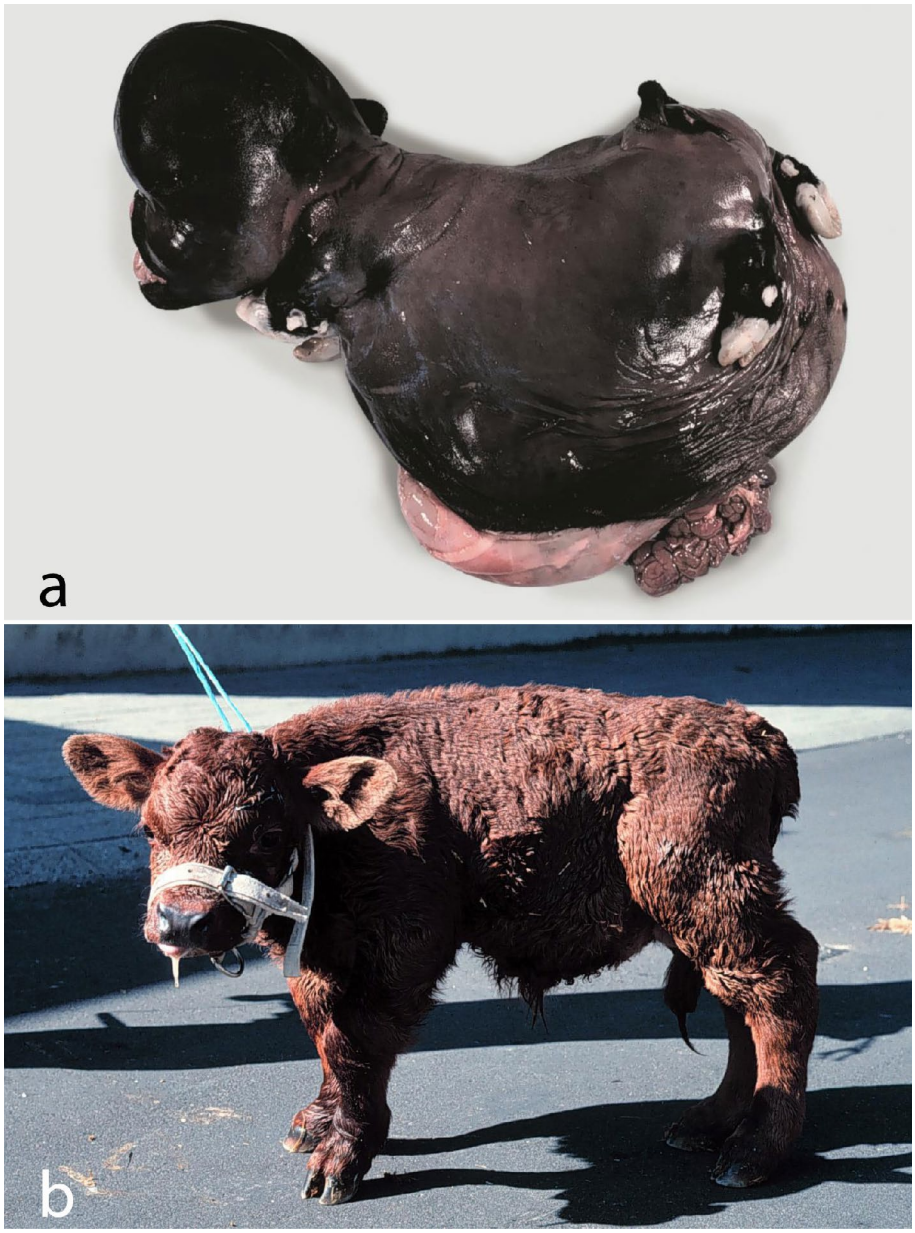


Fig. 5 Examples of the bulldog calf syndrome showing the wide spectrum of phenotypes. **a** The bulldog calf syndrome "prototype" as it occurs in Dexter cattle due to mutations in the *ACAN* gene. Aborted Dexter fetus; **b** Autosomal recessively inherited sublethal form of the bulldog calf syndrome as it occurred in Red Danish cattle [15]. The molecular basis of this phenotype is unknown

that also larger-sized genomic deletions cause a similar congenital phenotype and thereby expanding the knowledge on this condition by emphasizing that different mutations in *COL2A1* cause a uniform phenotype. For many genes it is known that the kind of genetic alteration influence the phenotypic outcome, e.g. the severity of a congenital defect varies or differs totally depending on the individual variant. Interestingly for *COL2A1* in cattle this seems not to be the case as different kinds of variants always cause an identical phenotype which is of importance for diagnostic pathologists. Furthermore, this report provides an overview of the phenotypic and allelic heterogeneity of the *COL2A1*-related BDS in cattle. This example highlights the utility of WGS-based precise diagnostics for understanding disorders linked to de novo mutations in animals with an available reference genome sequence

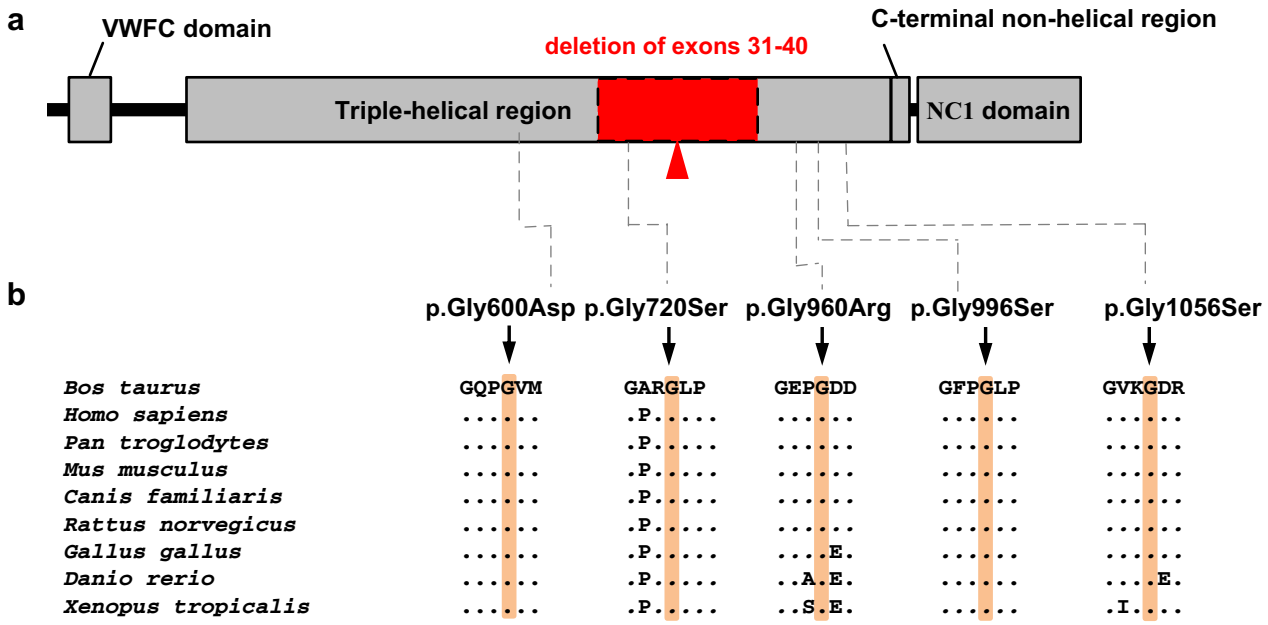


Fig. 6 Schematic diagram of the COL2A1 protein. **a** Domain and region information for the α1 chain of type II collagen (COL2A1) obtained from the UniProt database (http://www.uniprot.org/; accession number: P02459). Note that the large deletion (red arrow) affects the triple-helical region. **b** Multiple species alignment of the COL2A1 proteins across 9 different species showing a complete conservation of the affected residue. Note the pathogenic exchange of glycine residues of the Gly-x-y structural motif which is typical of triple-helical regions from collagen proteins reported in BDs-affected cases: p.Gly600Asp in Charolais Salers crossbreed calves, p.Gly720Ser in an isolated German Holstein calf [5], p.Gly960Arg in French Holstein calves [13], p.Gly996Ser in German Holstein calves [6] and p.Gly1056Ser in an isolated BD Holstein calf [14]. Protein sequences accession numbers in NCBI for each species are NP_001001135.2 (*Bos taurus*), NP_001835.3 (*Homo sapiens*), XP_509026.2 (*Pan troglodytes*), NP_112440.2 (*Mus musculus*), NP_001006952.1 (*Canis familiaris*), NP_037061.1 (*Rattus norvegicus*), NP_989757.1 (*Gallus gallus*), NP_571367.1 (*Danio rerio*) and NP_989220.1 (*Xenopus tropicalis*)

Table 1 Previously reported genetic variants of *COL2A1* causing the OMIA 001926-9913 bulldog calf syndrome in cattle

Inheritance	Type of variant	Variant[a]	Breed	References
AD, mosaicism	Missense	g.32307658G > A p.Gly960Arg	Holstein	[13]
AD, mosaicism	Splicing	g.32305226G > A	Holstein	[8]
AD, mosaicism	Missense	g.32301746G > A p.Gly600Asp	Charolais × Salers	[12]
AD, de novo	Missense	g.32303739G > A p.Gly720Ser	Holstein	[12]
AD, mosaicism	Missense	g.32308008G > A p.Gly996Ser	Holstein	[9]
AD, de novo	Missense	g.32308734G > A p.Gly1056Ser	Holstein	[14]

AD autosomal dominant, *OMIA* Online Mendelian Inheritance in Animals, https://omia.org/home/

[a] Given positions correspond to chromosome 5 of the ARS-UCD1.2 assembly and NP_001001135.2

and the need for continued surveillance for genetic disorders in cattle breeding. Genome sequencing might improve the precision of the clinicopathological diagnosis as sometimes unexpected variants in genes that were not known to be associated with a certain disorder could be detected.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13028-020-00548-w.

Additional file 1. EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 494 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available.

Additional file 2. List of the remaining private protein-coding variants after comparison of the genotypes of the calf with 494 cattle genome. A total of 20 private protein-changing single-nucleotide or short indel variants with a moderate or high predicted impact, located within 19 different genes or loci, were identified.

Abbreviations

BAM: Binary alignment map; BDS: Bulldog calf syndrome; Bp: Base pair; *COL2A1*: Collagen type II alpha 1 chain; EDTA: Ethylenediaminetetraacetic acid; Gly: Glycine; gnomAD: Genome Aggregation Database; GVCF: Genomic variant call format; IGV: Integrative Genomics Viewer; OMIA: Online Mendelian Inheritance in Animals, https://omia.org/home/; OMIM: Online Mendelian Inheritance in Man, https://omim.org/; PCR: Polymerase chain reaction; SNV: Single-nucleotide variant; WGS: Whole-genome sequencing.

Acknowledgements

The veterinary practice Kvægdyrlægerne Midt Aps, Bording, Denmark is thanked for submitting the calf and the owner for donating it. The authors would like to acknowledge the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome sequencing experiments and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure. The authors are grateful to Nathalie Besuchet-Schmutz for expert technical assistance.

Prior publication

Figure 5a has previously been published in Agerholm, JS. Inherited disorders in Danish cattle. APMIS. 2007;115 (suppl 122): 1–76.

Authors' contributions

JGPJ, AL and CD performed the genetic analyses. IMH carried out the bioinformatics. JSA performed the post-mortem examination. JGPJ drafted the manuscript and illustrations. JSA and CD designed the study, supervised the project and finalized the manuscript. All authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Funding

The study was funded by the Danish Bovine Genetic Disease Programme, internal funds for the University of Copenhagen, and the Swiss National Science Foundation.

Availability of data and materials

Whole-genome sequence data generated from the affected calf is available under study accession PRJEB18113 and sample accession SAMEA6528902 from the European Nucleotide Archive (ENA). In addition, further control genomes are listed in Additional file 1 and can also be accessed on ENA.

Ethics approval and consent to participate

This study did not require official or institutional ethical approval as it was not experimental.

Consent for publication

Not applicable.

Competing interests

JSA is editor-in-chief of Acta Veterinaria Scandinavia, but has not in any way been involved in or interacted with the journal's review process or editorial decision-making. The editor was blinded to the review process. The authors declare that they have no competing interests.

Author details

[1] Department of Veterinary Medical Sciences, University of Bologna, Via Tolara di Sopra 50, Ozzano dell'Emilia, 40064 Bologna, Italy. [2] Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstr. 109a, 3001 Bern, Switzerland. [3] Department of Veterinary Clinical Sciences, University of Copenhagen, Højbakkegaard Allé 5A, 2630 Taastrup, Denmark.

Received: 11 June 2020 Accepted: 30 August 2020
Published online: 07 September 2020

References

1. Agerholm JS, Arnbjerg J, Andersen O. Familial chondrodysplasia in Holstein calves. J Vet Diagn Invest. 2004;16:293–8.
2. Seligmann CG. Cretinism in calves. J Pathol. 1904;9:311–22.
3. Cavanagh JA, Tammen I, Windsor PA, Bateman JF, Savarirayan R, Nicholas FW, et al. Bulldog dwarfism in Dexter cattle is caused by mutations in *ACAN*. Mamm Genome. 2007;18:808–14.
4. Hayes BJ, Daetwyler HD. 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. Annu Rev Anim Biosci. 2019;7:89–102.
5. DePristo M, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
6. Cingolani P, Platts A, le Wang L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: sNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6:80–92.
7. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the Integrative Genomics Viewer. Cancer Res. 2017;77:31–4.
8. Agerholm JS, Menzi F, McEvoy FJ, Jagannathan V, Drögemüller C. Lethal chondrodysplasia in a family of Holstein cattle is associated with a de novo splice site variant of *COL2A1*. BMC Vet Res. 2016;12:100.
9. Reinartz S, Mohwinkel H, Sürie C, Hellige M, Feige K, Eikelberg D, et al. Germline mutation within *COL2A1* associated with lethal chondrodysplasia in a polled Holstein family. BMC Genomics. 2017;18:762.
10. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581:434–43.
11. Shoulders MD, Raines RT. Collagen structure and stability. Annu Rev Biochem. 2009;78:929–58.
12. Bourneuf E, Otz P, Pausch H, Jagannathan V, Michot P, Grohs C, et al. Rapid discovery of de novo deleterious mutations in cattle enhances the value of livestock as model species. Sci Rep. 2017;7:11466.
13. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat Genet. 2014;46:858–65.
14. Häfliger IM, Behn H, Freick M, Jagannathan V, Drögemüller C. A *COL2A1 de novo* variant in a Holstein bulldog calf. Anim Genet. 2019;50:113–4.
15. Andresen E, Christensen K, Jensen PT, Venge O, Rasmussen PG, Basse A, et al. Dværgvækst hos kvæg af Rød Dansk Malkerace (Achondroplasia bovis typus RDM). Nord VetMed. 1974;26:681–91.

Publisher's Note

A *de novo* mutation in *COL1A1* in a Holstein calf with osteogenesis imperfecta type II

Journal:	Animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review
Displayed version:	published version
DOI:	10.3390/ani11020561

animals

MDPI

Article

A De Novo Mutation in *COL1A1* in a Holstein Calf with Osteogenesis Imperfecta Type II

Joana G. P. Jacinto [1,2], Irene M. Häfliger [2], Fintan J. McEvoy [3], Cord Drögemüller [2,*,†] and Jørgen S. Agerholm [3,†]

[1] Department of Veterinary Medical Sciences, University of Bologna, 40064 Ozzano Emilia, Italy; joana.goncalves2@studio.unibo.it
[2] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch
[3] Department of Veterinary Clinical Sciences, University of Copenhagen, Dyrlægevej 16, DK 1870 Copenhagen, Denmark; fme@sund.ku.dk (F.J.M.); jager@sund.ku.dk (J.S.A.)
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-31-631-2529
† These authors contributed equally to this work.

Simple Summary: Skeletal connective tissue disorders represent a heterogeneous group of inherited disorders mostly monogenically inherited. Heritable connective tissue disorders such as osteogenesis imperfecta (OI) belong to this group. Herein, an affected calf showing congenital bone lesions such as intrauterine fractures, abnormally shaped long bones and localized arthrogryposis resembling OI type II is reported. Whole-genome sequencing (WGS) identified a most likely disease-causing mutation in the *COL1A1* gene. The *COL1A1* gene is known to be associated with dominant inherited OI type II forms in humans and sporadically in dogs and cattle, but so far, a variant in the fibrillar collagen NC1 domain has not been shown to cause a similar phenotype in domestic animals. We assume that the herein identified most-likely causative variant occurred either within the parental germlines or post-zygotically in the developing embryo. Rare lethal disorders such as OI in livestock are usually not diagnosed to the molecular level, mainly because of the lack of resources and diagnostic tools. WGS-based precision diagnostics allows understanding rare disorders and supports the value of surveillance of cattle breeding populations for harmful genetic disorders.

Abstract: Osteogenesis imperfecta (OI) type II is a genetic connective tissue disorder characterized by bone fragility, severe skeletal deformities and shortened limbs. OI usually causes perinatal death of affected individuals. OI type II diagnosis in humans is established by the identification of heterozygous mutations in genes coding for collagens. The purpose of this study was to characterize the pathological phenotype of an OI type II-affected neonatal Holstein calf and to identify the causative genetic variant by whole-genome sequencing (WGS). The calf had acute as well as intrauterine fractures, abnormally shaped long bones and localized arthrogryposis. Genetic analysis revealed a private heterozygous missense variant in *COL1A1* (c.3917T>A) located in the fibrillar collagen NC1 domain (p.Val1306Glu) that most likely occurred de novo. This confirmed the diagnosis of OI type II and represents the first report of a pathogenic variant in the fibrillar collagen NC domain of *COL1A1* associated to OI type II in domestic animals. Furthermore, this study highlights the utility of WGS-based precise diagnostics for understanding congenital disorders in cattle and the need for continued surveillance for rare lethal genetic disorders in cattle.

Keywords: cattle; *Bos taurus*; collagenopathy; skeletal disorder; bone disease; rare diseases; precision medicine; whole-genome sequencing

Citation: Jacinto, J.G.P.; Häfliger, I.M.; McEvoy, F.J.; Drögemüller, C.; Agerholm, J.S. A De Novo Mutation in *COL1A1* in a Holstein Calf with Osteogenesis Imperfecta Type II. *Animals* **2021**, *11*, 561. https://doi.org/10.3390/ani11020561

Academic Editors: Francesca Ciotola, Sara Albarella and Vincenzo Peretti

Received: 29 January 2021
Accepted: 19 February 2021
Published: 20 February 2021

1. Introduction

Osteogenesis imperfecta (OI) encompasses a heterogeneous group of rare genetic connective tissue disorders characterized by skeletal abnormalities, leading to bone fragility,

deformity, low bone mass and growth deficiency [1,2]. Decrease bone strength predisposes to low-trauma factures or factures in atypical regions [3]. In humans, extra-skeletal manifestations of OI may include joint hypermobility, dentinogenesis imperfecta, blue sclera, hearing loss, and more rarely pulmonary and cardiovascular complications, and muscle weakness [2].

Currently, OI in humans is classified into as many as 18 types; the classification depends on the genetic causes, severity and clinical observation [4]. OI types I-IV are mainly associated with autosomal dominant variants in *COL1A1* and *COL1A2*; OI type V is a less frequent dominantly inherited form associated with variants in the novel gene *IFTM5*. The remaining types of OI, which usually arise at much lower frequency are autosomal recessive diseases while OI type XVIII has an X-linked inheritance pattern [2,4].

Since the second half of the 20th century several forms of OI have been reported in domestic animal species, including sheep [5], cats [6], dogs [7] and cattle [8]. Until now, one OI-related causative dominant variant in the *COL1A1* is known in dogs (OMIA 002126-9615) [9] and two in cattle (OMIA 002127-9913) [10,11]. In dogs, three OI-related causative dominant variants in the *COL1A2* are also known (OMIA 002112-9615) [12–14] and also in this species, an OI-related recessive form has been associated with a missense variant in *SERPINH1* (OMIA 001483-9615) [15].

In this study, we aimed to characterize an OI-affected Holstein calf, and to identify the causative genetic variant associated with the disorder using whole-genome sequencing (WGS).

2. Materials and Methods

2.1. Pathological Investigation

A male Holstein calf with a weight of 30.6 kg was delivered at gestation day 264 (normal gestation 281 days (mean)). The pregnancy was the result of insemination with semen of a purebred Holstein sire on a Holstein dam. The parents were not related within at least four generations. The calf was immediately humanely euthanized by an intravenous overdose of barbiturate upon delivery due to severe malformations. The carcass was submitted for necropsy during which radiographs were taken of the limbs. Tissue samples were taken during the necropsy and included internal organs, brain, metacarpus, metatarsus, humerus and tibia. These were fixed in 10 % neutral buffered formalin. Bone specimens were thereafter transferred to an aqueous solution containing sodium formate (0.5 mol/L) and formic acid (0.5 mol/L) (Kristensen's decalcifying medium) until suitable for cutting. The tissues were thereafter trimmed, processed by routine methods, paraffin embedded, sectioned at 2 µm and stained with haematoxylin and eosin.

2.2. DNA Samples

Genomic DNA was extracted from skin and cartilage taken from the ear of the calf, from EDTA blood of its dam and from semen of its sire using Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland).

2.3. Whole-Genome Sequencing

Using genomic DNA of the affected calf an individual PCR-free fragment library with approximately 400 bp inserts was prepared and sequenced for 150 bp paired-end reads using the NovaSeq6000 system (Illumina, San Diego, CA, USA). The sequenced reads were mapped to the ARS-UCD1.2 reference genome resulting in an average read depth of approximately 18.1× [16], and single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format (GVCF) files were in accordance with the 1000 Bull Genomes Project processing guidelines of run 7 (www.1000bullgenomes.com (accessed on 31 August 2018)) [17], except for the trimming, which was performed using fastp [18]. Further preparation of the genomic data had been done according to Häfliger et al. 2020 [19]. The impact of the called variants was functionally annotated with snpeff v4.3 [20], using

the NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/ (accessed on 31 August 2018)), which resulted in the final VCF file, including all individual variants and their functional annotations. In order to find private variants, we compared the genotypes of the affected calf with 496 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that are publicly available (Table S1) in the European Nucleotide Archive (SAMEA6528897 is the sample accession number of the affected calf; http://www.ebi.ac.uk/en (accessed on 7 February 2020)). Integrative Genomics Viewer (IGV) [21] software version 2.0 was used for visual inspection of genome regions containing possible candidate genes.

2.4. Targeted Genotyping

Polymerase chain reaction (PCR) and Sanger sequencing were used to validate and genotype the variant identified from WGS. The *COL1A1* missense variant (NM_001034039.2: g.36473359T>A) was genotyped using the following primers: 5′- ATCTTACTTTGCCCCACCCC-3′ (forward primer) and 5′-GGCTACAAGGTCCAG CTCAC-3′ (reverse primer). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA).

2.5. Evaluation of the Molecular Consequences of Amino Acid Substitutions

PROVEAN [22] and DynaMut [23] were used to predict the functional consequences of the discovered variant on protein. For multispecies sequence alignments the following NCBI proteins accessions were used: NP_001029211.1 (*Bos taurus*), NP_000079.2 (*Homo sapiens*), XP_001169409.1 (*Pan troglodytes*), XP_001096194.2 (*Macaca mulatta*), NP_001003090.1 (*Canis lupus*), NP_031768.2 (*Mus musculus*), NP_445756.1 (*Rattus norvegicus*), NP_954684.1 (*Danio rerio*), NP_001011005.1 (*Xenopus tropicalis*).

2.6. Sequence Accessions

All references to bovine *COL1A1* gene correspond to the NCBI accessions NM_001034039.2 (*COL1A1* mRNA) and NP_001029211.1 (COL1A1 protein). For the protein, structure of COL1A1 the Uniprot database (https://www.uniprot.org/ (accessed on 31 August 2018)) accession number P02453 was used.

3. Results

3.1. Pathological Phenotype

The limbs appeared shortened with bilateral symmetric flexion of the fetlock joint and mild lateral rotation of the digits. Flexion of the thoracic limb metacarpo-phalangeal joints was mild (15°) while it was almost 90° in the pelvic limbs. The tibiotarsal joints were extended (Figure 1a). Multiple long bones were abnormal and especially the metacarpal and metatarsal bones were bowed and appeared of reduced diameter (Figure 1b). Several transverse or oblique fractures, some with displaced fracture ends, were present in the long bones and the left hemimandible. Some fractures had fibrous callous formation and cortical bone proliferation. Non-aligned fracture ends showed abnormal healing (Figure 1c,d), while others were acute. Teeth appeared normal as did the color of the sclera.

Histology only revealed lesions in the bones. The epiphyseal trabeculae were reduced in size and number (Figure 2a). Also, ossification was reduced as chondroid matrix was widely present in bone spicules and occasional islets of chondroid matrix were seen. The epiphyseal growth lines were normal, but as for the epiphyses, the metaphyseal and diaphyseal trabecular bone was of reduced amount and quality. The cortical bone appeared thinner (Figure 2b). The bony ends of the fractured left tibia were completely covered by a prominent fibrous callus, while the fracture of the left metacarpus had no signs of repair.

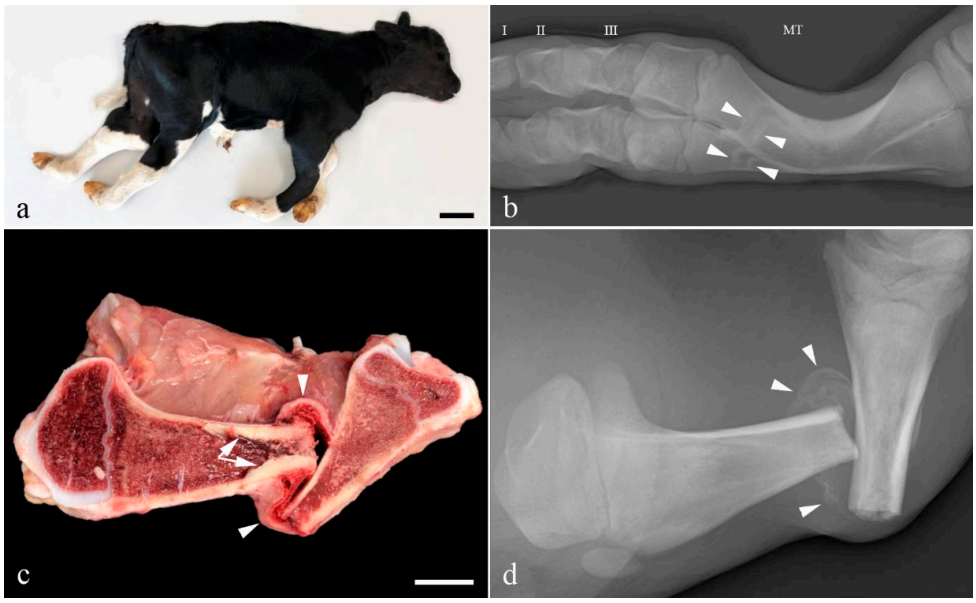


Figure 1. Gross morphology of the OI affected Holstein calf. (**a**) The limbs appear shortened and the distal joints are flexed. Bar = 10 cm; (**b**) Radiograph of a metatarsal bone (MT). The diaphysis is curved and transverse lines of sclerosis (arrow heads) are present distally. These may represent sites of healed fracture or growth arrest lines. Phalanx I, II and III are indicated by their respective numbers; (**c**) Longitudinal section through the left tibia. A transverse fracture with dislocation of the fracture ends is seen. Prominent osseous endostal proliferation has developed in the proximal part of the fracture (arrows). The fracture is surrounded by fibrosis (arrow heads). The width of the cortex is un-uniform with the caudal part of the proximal diaphysis/metaphysis being thin. Bar = 2.5 cm; (**d**) Radiograph of the specimen shown in (**c**) before longitudinal sectioning. The non-aligned fracture ends are surrounded by partly mineralized fibrous callus (arrow heads).

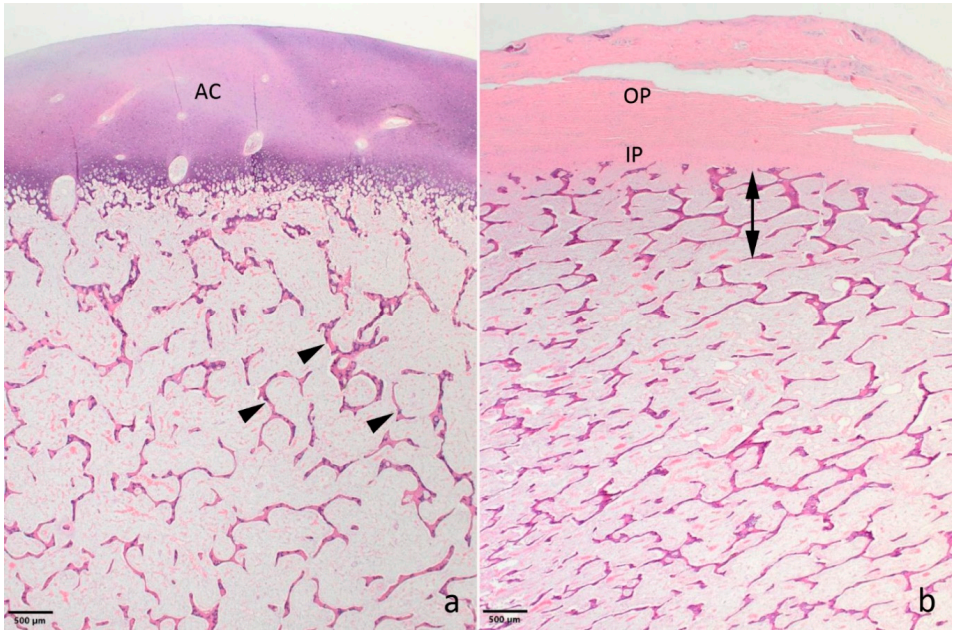


Figure 2. Photomicrographs showing the bone morphology of the OI affected Holstein calf. (**a**) The epiphyseal spongiosa is characterized by small trabeculae with remnants of chondroid matrix (arrow heads). The amount of spongious bone is reduced resulting in wide spaces between the trabeculae. AC: articular cartilage. Distal epiphysis, metacarpus. (**b**) The compacta (double headed arrow) is poorly developed and difficult to distinguish from the spongious bone. As for the epiphyses, the trabecular bone is poorly developed. IP and OP: inner and outer layer of the periost, respectively. Metaphysis, distal metatarsus. (**a**,**b**): Haematoxylin and eosin.

3.2. Genetic Analysis

Filtering of WGS for private variants present in the affected calf and absent in the 496 available control genomes, identified 14 protein-changing variants with a predicted high and moderate impact on the encoded protein. They were found to be heterozygous exclusively in the OI-affected calf and absent in the 496 control genomes that were sequenced in the course of other ongoing projects at the Institute of Genetics. These variants were further investigated for their occurrence in a global control cohort of 3103 genomes of a variety of breeds 1000 Bull Genomes Project run 7 [17], which revealed five protein-changing variants exclusively present heterozygous in the genome of the affected calf (Table S2).

These five variants were subsequently visually inspected using IGV software confirming all as true variants. Of all the remaining private variants, only one occurred in an obvious candidate gene for OI (Figure 3a). The heterozygous variant at chr19:36473359T>A represents a missense variant in COL1A1 (NM_001034039.2: c.3917T>A; Figure 3c). This variant alters the encoded amino acid of COL1A1 residue 1306 (XP_024835395.1: p.Val1306Glu) located in the fibrillar collagen NC1 domain (Figure 3d,e). Furthermore, the valine to glutamine substitution affects an evolutionary conserved amino acid (Figure 3d) and was predicted to be deleterious (PROVEAN score −4.96) and destabilizing (DynaMut, ΔΔG: −0.127 kcal/mol). To confirm and evaluate the presence of the COL1A1 variant, the affected genomic region was amplified by PCR and Sanger sequenced in the calf, its sire and dam (Figure 3b). Analyzing the sequencing data, we observed that the calf was indeed heterozygous for the detected COL1A1 variant whereas the sire and dam were both homozygous for the wild type allele (Figure 3b). This showed that the mutation most likley arose spontaneously in the affected calf and finally confirmed the diagnosis of OI type II.

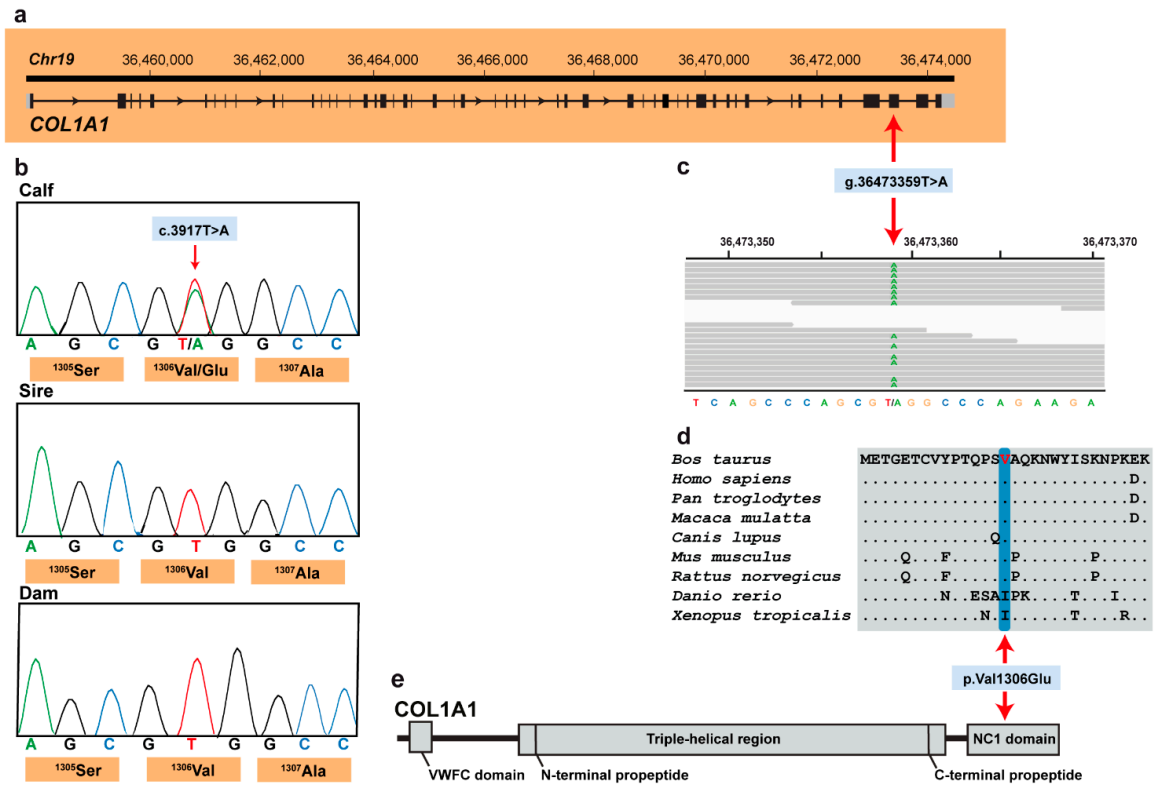


Figure 3. *COL1A1* missense variant in an OI type II-affected Holstein calf. (**a**) *COL1A1* gene structure showing the variant location on chromosome 19, exon 49 (red arrow); (**b**) Electropherograms showing the heterozygous genotype of the calf, and the absence of the variant in its dam genome and in the germline of its sire. (**c**) IGV screenshot presenting the Chr19: g.36473359T>A variant in the affected calf. (**d**) Multiple sequence alignment of the collagen alpha-1(I) chain of the COL1A1 protein encompassing the region of the p.Val1306Glu variant demonstrates complete evolutionary conservation across species. (**e**) Schematic representation of COL1A1 protein and its three functional domains.

4. Discussion

The identified missense variant in *COL1A1* in an obvious candidate gene represents the most likely pathogenic variant associated with the observed OI phenotype. As for humans, OI in cattle is a disorder characterized by bone fragility with perinatal fractures, severe bowing of long bones and reduced mineralization. Furthermore, the OI type II form is frequently lethal in utero or shortly after birth due to severe bone fragility and respiratory insufficiency. In cattle, forms of OI type II (OMIA 002127-9913) have been reported in Fleckvieh and Red Angus cattle associated to dominant acting pathogenic variants in *COL1A1* inherited from mosaic sires [10,11]. In humans, this disorder (OMIM 166210) is linked to pathogenic variants in *COL1A1* and *COL1A2* with a dominant pattern of inheritance. Thus, cases of OI with lesions typical for type II OI could be suspected of having a defect in *COL1A1* or *COL1A2* genes; a suspicion that is helpful when analysing WGS data. In this case, filtering for private variants lead to the identification of a missense variant in *COL1A1*. We assume that this spontaneous mutation most likely occurred either within the parental germlines or post-zygotically in the developing embryo. The mutant allele was detected neither in the dam nor in the sire, given that the variant was not found in the paternal germ line DNA which was analysed. This means that the missense variant was exclusively present in heterozygous state in the affected offspring only. Therefore, it appears more plausible that the identified mutation arose indeed de novo spontaneously in the very early development of the calf. However, a low level mosaicism in the dam cannot be excluded given that the DNA which was analysed was not from the germline.

The identified deleterious variant and the conservation of the affected amino acid residue of COL1A at the position 1306 suggest that this variant is most likely pathogenic. The affected valine residue is conserved across mammals and corresponds to isoleucine in more distant related species such as clawed frogs or zebrafish. As isoleucine and valine are highly similar amino acids having large aliphatic hydrophobic side chains their molecules are rigid, and their mutual hydrophobic interactions are important for the correct folding of proteins, as these chains tend to be located inside of the protein molecule. Because glutamic acid expressed by the mutant allele carries a long hydrophilic acidic group with strong negative charge it most likely impairs proper folding. Type I collagen is a member of group I collagen (fibrillar forming collagen) and is located in the extracellular matrix (ECM). There is a wide range of connective tissue disorders that occur from genetic abnormalities in ECM proteins as for example OI. In contrast to heterozygous null mutations, most of the more severe ECM disorders are caused by heterozygous missense mutations that interrupt the protein structure [24]. Variants in genes coding for collagens provide the perfect scenario for the impact of dominant negative protein structural mutations. In addition to variants in *COL1A1* and *COL1A2* associated with OI, many pathogenic variants in *COL2A1* and *COL10A1* are linked to chondrodysplasias [24].

Interestingly, the identified missense variant in this study occurred in the fibrillar collagen NC1 domain, and by analogy with substitutions in this domain, as in collagen alpha-10(X) in Schmid metaphysical chondrodysplasias (OMIM 156500), it would be expected to disrupt the collagen homotrimer assembly and secretion leading to proteosomal degradation of the unassembled collagen chains [25–28]. To the best of our knowledge, no pathogenic variant in fibrillar collagen NC1 domain of COL1A1 associated with OI has been reported in domestic animal species and therefore represents the first large animal model for mutations occurring in this domain of COL1A1.

5. Conclusions

This is the first report of a most likely pathogenic missense variant in OI type II-affected cattle affecting the fibrillar collagen NC domain of COL1A1. So far, disease-causing variants in *COL1A1* in domestic animals were found only in the triple-helical region of the encoded collagen protein. This case demonstrates that pathogenic variants in the other domains of the COL1A1 protein also cause a similar congenital phenotype. Therefore, this finding expands the spectrum of *COL1A1* mutations that cause a uniform phenotype. Furthermore,

this example highlights the utility of WGS-precise diagnosis for understanding sporadic cases of congenital disorders associated to de novo mutations and the need for continued surveillance of genetic lethal disorders in cattle breeding.

Supplementary Materials: Supplementary materials can be found at https://www.mdpi.com/2076-2615/11/2/561/s1. Table S1: EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 494 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available, Table S2: List of the remaining variants after the comparison to the global control cohort of 3103 genomes of other breeds (1000 Bull Genomes Project run 7; www.1000bullgenomes.com (accessed on 31 August 2018)), revealing 5 protein-changing variants only present in the genome of the OI type II-affected calf. These 5 variants with a moderate predicted impact on the encoded protein were located within 5 different genes or loci. Note that the predicted pathogenic variant NM_001034039.2: c.3917T>A is the only one located in a candidate gene for bone diseases.

Author Contributions: Conceptualization, C.D. and J.S.A.; methodology, I.M.H., F.J.M., J.S.A. and C.D.; software, I.M.H.; validation, J.G.P.J.; formal analysis, J.G.P.J., I.M.H., F.J.M., J.S.A. and C.D.; investigation, J.G.P.J. and J.S.A.; resources, C.D. and J.S.A.; data curation, I.M.H.; writing—original draft preparation, J.G.P.J.; writing—review and editing, J.G.P.J., C.D., F.J.M. and J.S.A.; visualization, J.S.A. and J.G.P.J.; supervision, C.D. and J.S.A.; project administration, C.D. and J.S.A.; funding acquisition, C.D. and J.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: The post-mortem investigation was performed as part of the Danish surveillance program for genetic diseases in cattle and funded jointly by SEGES and the University of Copenhagen. This study was partially funded by the Swiss National Science Foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but part of a clinical and pathological veterinary diagnostic case.

Data Availability Statement: The whole-genome data of the affected calf is freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA6528897.

Acknowledgments: Tine Skau, Vendsyssel Landdyrlæger is acknowledged for referring the case and the owner for donating the calf for research. We thank Nathalie Besuchet-Schmutz for expert technical assistance and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Marini, J.C.; Forlino, A.; Cabral, W.A.; Barnes, A.M.; Antonio, J.D.S.; Milgrom, S.; Hyland, J.C.; Körkkö, J.; Prockop, D.J.; De Paepe, A.; et al. Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: Regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Hum. Mutat.* **2007**, *28*, 209–221. [CrossRef] [PubMed]
2. Rossi, V.; Lee, B.; Marom, R. Osteogenesis imperfecta: Advancements in genetics and treatment. *Curr. Opin. Pediatr.* **2019**, *31*, 708–715. [CrossRef] [PubMed]
3. Peddada, K.V.; Sullivan, B.T.; Margalit, A.; Sponseller, P.D. Fracture Patterns Differ Between Osteogenesis Imperfecta and Routine Pediatric Fractures. *J. Pediatr. Orthop.* **2018**, *38*, e207–e212. [CrossRef] [PubMed]
4. Marini, J.C.; Cabral, W.A. *Osteogenesis Imperfecta Genetics of Bone Biology and Skeletal Disease*; Elsevier Inc: Amsterdam, The Netherlands, 2018.
5. Kater, J.C.; Hartley, W.; Dysart, T.; Campbell, A. Osteogenesis imperfecta and bone resorption: Two unusual skeletal abnormalities in young lambs. *New Zealand Veter- J.* **1963**, *11*, 41–46. [CrossRef]
6. Cohn, L.; Meuten, D.J. Bone fragility in a kitten: An osteogenesis imperfecta-like syndrome. *J. Am. Veter- Med. Assoc.* **1990**, *197*, 98–100.
7. Calkins, E.; Kalm, D.; Diner, W.C. Idiopathic familial osteoporosis in dogs: "Osteogenesis imperfecta". *Ann. N. Y. Acad. Sci.* **1956**, *64*, 410–423. [CrossRef]
8. Jensen, P.T.; Rusmussen, P.G.; Basse, A. Congenital osteogenesis imperfecta in Charolais cattle. *Nord. Vet. Med.* **1976**, *28*, 304–308.
9. Campbell, B.G.; Wootton, J.A.; MacLeod, J.N.; Minor, R.R. Sequence of Normal Canine COL1A1 cDNA and Identification of a Heterozygous α1(I) Collagen Gly208AIa Mutation in a Severe Case of Canine Osteogenesis Imperfecta. *Arch. Biochem. Biophys.* **2000**, *384*, 37–46. [CrossRef]

10. Petersen, J.L.; Tietze, S.M.; Burrack, R.M.; Steffen, D.J. Evidence for a de novo, dominant germ-line mutation causative of osteogenesis imperfecta in two Red Angus calves. *Mamm. Genome* **2019**, *30*, 81–87. [CrossRef] [PubMed]
11. Bourneuf, E.; Otz, P.; Pausch, H.; Jagannathan, V.; Michot, P.; Grohs, C.; Piton, G.; Ammermüller, S.; Deloche, M.C.; Fritz, S.; et al. Rapid Discovery of De Novo Deleterious Mutations in Cattle Enhances the Value of Livestock as Model Species. *Sci. Rep.* **2017**, *7*, 1–19. [CrossRef] [PubMed]
12. Letko, A.; Zdora, I.; Hitzler, V.; Jagannathan, V.; Beineke, A.; Möhrke, C.; Drögemüller, C. A de novo in-frame duplication in the COL 1A2 gene in a Lagotto Romagnolo dog with osteogenesis imperfecta. *Anim. Genet.* **2019**, *50*, 786–787. [CrossRef] [PubMed]
13. Campbell, B.G.; Wootton, J.A.M.; MacLeod, J.N.; Minor, R.R. Canine COL1A2 Mutation Resulting in C-Terminal Truncation of Pro-α2(I) and Severe Osteogenesis Imperfecta. *J. Bone Miner. Res.* **2001**, *16*, 1147–1153. [CrossRef] [PubMed]
14. Quist, E.M.; Doan, R.; Pool, R.R.; Porter, B.F.; Bannasch, D.L.; Dindot, S.V. Identification of a Candidate Mutation in the COL1A2 Gene of a Chow Chow With Osteogenesis Imperfecta. *J. Hered.* **2017**, *109*, 308–314. [CrossRef] [PubMed]
15. Drögemüller, C.; Becker, D.; Brunner, A.; Haase, B.; Kircher, P.; Seeliger, F.; Fehr, M.; Baumann, U.; Lindblad-Toh, K.; Leeb, T. A Missense Mutation in the SERPINH1 Gene in Dachshunds with Osteogenesis Imperfecta. *PLoS Genet.* **2009**, *5*, e1000579. [CrossRef]
16. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience* **2020**, *9*, 1–9. [CrossRef]
17. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef] [PubMed]
18. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinform.* **2018**, *34*, i884–i890. [CrossRef]
19. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
20. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]
21. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant Review with the Integrative Genomics Viewer. *Cancer Res.* **2017**, *77*, e31–e34. [CrossRef] [PubMed]
22. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinform.* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]
23. Rodrigues, C.H.M.; Pires, D.E.V.; Ascher, D.B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* **2018**, *46*, W350–W355. [CrossRef]
24. Lamandé, S.R.; Bateman, J.F. Genetic Disorders of the Extracellular Matrix. *Anat. Rec. Adv. Integr. Anat. Evol. Biol.* **2019**, *303*, 1527–1542. [CrossRef] [PubMed]
25. Kong, L.; Shi, L.; Wang, W.; Zuo, R.; Wang, M.; Kang, Q. Identification of two novel COL10A1 heterozygous mutations in two Chinese pedigrees with Schmid-type metaphyseal chondrodysplasia. *BMC Med. Genet.* **2019**, *20*, 1–9. [CrossRef]
26. Chan, D.; Cole, W.G.; Rogers, J.; Bateman, J.F. A mutation in the conserved NC1 domain of type X collagen prevents in vitro multimer assembly resulting in a Schmid-type metaphyseal chondrodysplasia. *Matrix Biol.* **1994**, *14*, 396–397. [CrossRef]
27. Chan, D.; Weng, Y.M.; Golub, S.; Bateman, J.F. Type X collagen NC1 mutations produced by site-directed mutagenesis prevent in vitro assembly. *Ann. N. Y. Acad. Sci.* **1996**, *785*, 231–233. [CrossRef] [PubMed]
28. Wilson, R.; Freddi, S.; Bateman, J.F. Collagen X Chains Harboring Schmid Metaphyseal Chondrodysplasia NC1 Domain Mutations Are Selectively Retained and Degraded in Stably Transfected Cells. *J. Biol. Chem.* **2002**, *277*, 12516–12524. [CrossRef] [PubMed]

Trisomy 29 in a stillborn Swiss Original Braunvieh calf

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.12929

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

BRIEF NOTES

doi: 10.1111/age.12929

Trisomy 29 in a stillborn Swiss Original Braunvieh calf

Irene Monika Häfliger*, Franz Seefried† and Cord Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland; †Qualitas AG, Zug 6300, Switzerland

Accepted for publication 01 March 2020

Background

Calves being stillborn is mostly reported to be due to infectious diseases, and the calves do not usually show obvious malformations.[1] However, congenital disorders owing to monogenic variation have been described, leading to a variety of severe skeletal malformations, e.g. chondrodysplasia (OMIA 000187-9913), including bulldog calf syndrome (OMIA 001926-9913) and arachnomelia (OMIA 000059-9913; OMIA 001541-9913). In addition, chromosomal disorders, like aneuploidies, defined as the copy number variation of whole chromosomes, sporadically explain severely malformed stillborns, as recently shown in pulmonary hypoplasia with anasarca syndrome (OMIA 001562-9913).[2]

Own analysis

A 4-year old Original Braunvieh cow gave birth after 271 days of gestation to a stillborn malformed calf, reported as having dwarfism with severe facial anomalies (Fig. S1a). The purebred female calf apparently showed a proportional dwarfism phenotype with no obvious limb abnormalities. However, the head of the stillborn calf, resembling hydrocephalus, appeared to have severe facial dysplasia and brachygnathia, where the lower jaw appears to be significantly shortened (Fig. S1a). Unfortunately, further examinations, including necropsy, were not performed. It was the second gestation of the dam. The confirmed father was used as a natural service sire on 16 different farms and had produced 78 (35 female and 43 male) normal offspring. Nevertheless, owing to the severity of the malformations a genetic origin was suspected, possibly owing to a *de novo* mutation event. Whole-genome sequencing of the affected calf was performed using genomic DNA obtained from a submitted ear tissue sample. The obtained sequence reads were mapped to the latest reference genome ARS-UCD1.2, as described before.[2] This resulted in an average genome coverage of 17×. The raw data are available on the European Nucleotide Archive under project accession no. PRJEB18113 and sample accession SAMEA6528888. Initially, the read depth across chromosomes was analyzed in order to evaluate large duplications and deletions. Thereby, we detected an increased read depth across the whole of chromosome 29, compared with the rest of the genome (Fig. S1b). The bamqc-file of the qualimap[3] program reports a read depth ranging between 15.3× and 16.3× for autosomes 1–28, whereas chromosome 29 shows a 1.5 times higher coverage of 23.9×, which indicates a trisomy of bovine chromosome 29.

Comments

We postulate that the detected extra copy of chromosome 29 most likely caused the disorder in the presented case. The literature describes few cases of trisomy observed in cattle. Among other anomalies, all of the reported cases in cattle describe facial dysplasia, brachygnatia and/or palatoschisis, as well as malformations of the limbs.[2,4-10] Whereas some trisomies such as 18, 20 and 21 are characterized as lethal, living animals with trisomy 28 have been described.[7] Trisomies of chromosome 22 are described by a variety of phenotypes, ranging from viable calves to highly malformed aborts.[4,8,9] Most trisomies, as they have never been observed, are suspected to be lethal and thereby cause spontaneous abortions at different stages of gestation.[9] The case described herein seems to be the first evidence of a nearly full-term delivered calf being born with trisomy in bovine chromosome 29. More common chromosomal anomalies in cattle are incidences of translocations. In particular, in the so-called Robertsonian translocation chromosome 29 fuses to chromosome 1.[11] However, as the increased read depth in chromosome 29 is clearly indicating a trisomy, we exclude the possibility of a translocation.

In conclusion, we suspect that the trisomy of bovine chromosome 29 led to the lethal phenotype, including facial dysplasia in the stillborn calf. The origin of bovine trisomy 29 must have been due to non-disjunction during gametogenesis in one of the parents or during early cell division in the fertilized egg.

Acknowledgements: We thank the farmer who sent us the samples and took pictures of the presented case. We acknowledge the assistance of the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome re-sequencing experiments. Furthermore, we thank the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computing infrastructure. This study was funded by the Swiss National Science Foundation.

References

1 Wolf-Jackel G.A. *et al.* (2020) *Acta Vet Scand* **62**, 1.
2 Häfliger I.M. *et al.* (2020) *Anim Genet* [Eprint ahead: https://doi.org/10.1111/age.12920].
3 Okonechnikov K. *et al.* (2016) *Bioinformatics* **32**, 292–4.
4 Agerholm J.S. & Christensen K. (1993) *J Vet Med* **40**, 576–81.
5 Gallagher DS, *et al.* (1999) *Vet Path* **36**, 448–51.
6 Herzog A, *et al.* (1977) *Ann Genet Sel Anim* **9**, 471–91.
7 Iannuzzi L. *et al.* (2001) *Hereditas* **134**, 147–51.

8 Mayr B. *et al.* (1985) *Cytogenet Cell Genet* **39**, 77–9.
9 Schmutz S.M. *et al.* (1996) *J Vet Diagn Invest* **8**, 91–5.
10 Lioi M.B. *et al.* (1995) *Genet Sel Evol* **27**, 473–6.
11 De Lorenzi L. *et al.* (2012) *J Anim Breed Genet* **129**, 409–16.

Correspondence: C. Drögemüller
(cord.droegemueller@vetsuisse.unibe.ch)

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 (a) Gross morphology of the stillborn Original Braunvieh calf with ~50 cm body length. Note the round-shaped hydrocephalic-like head with extreme brachygnathia. (b) Genome-wide average read depth analysis (500 kb window size). Note the 1.5-fold increase in coverage of the entire chromosome 29 (red arrow).

KCNG1-related syndromic form of congenital neuromuscular channelopathy in a crossbred calf

Journal:	Genes
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.3390/genes12111792

genes

MDPI

Article

KCNG1-Related Syndromic Form of Congenital Neuromuscular Channelopathy in a Crossbred Calf

Joana G. P. Jacinto [1,2], Irene M. Häfliger [2], Eylem Emek Akyürek [3], Roberta Sacchetto [3], Cinzia Benazzi [1], Arcangelo Gentile [1] and Cord Drögemüller [2,*]

[1] Department of Veterinary Medical Sciences, University of Bologna, 40064 Ozzano Emilia, Italy; joana.goncalves2@studio.unibo.it (J.G.P.J.); cinzia.benazzi@unibo.it (C.B.); arcangelo.gentile@unibo.it (A.G.)
[2] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch
[3] Department of Comparative Biomedicine and Food Science, University of Padova, 35020 Legnaro, Italy; eylememek.akyurek@phd.unipd.it (E.E.A.); roberta.sacchetto@unipd.it (R.S.)
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch

Abstract: Inherited channelopathies are a clinically and heritably heterogeneous group of disorders that result from ion channel dysfunction. The aim of this study was to characterize the clinicopathologic features of a Belgian Blue x Holstein crossbred calf with paradoxical myotonia congenita, craniofacial dysmorphism, and myelodysplasia, and to identify the most likely genetic etiology. The calf displayed episodes of exercise-induced generalized myotonic muscle stiffness accompanied by increase in serum potassium. It also showed slight flattening of the splanchnocranium with deviation to the right side. On gross pathology, myelodysplasia (hydrosyringomielia and segmental hypoplasia) in the lumbosacral intumescence region was noticed. Histopathology of the muscle profile revealed loss of the main shape in 5.3% of muscle fibers. Whole-genome sequencing revealed a heterozygous missense variant in *KCNG1* affecting an evolutionary conserved residue (p.Trp416Cys). The mutation was predicted to be deleterious and to alter the pore helix of the ion transport domain of the transmembrane protein. The identified variant was present only in the affected calf and not seen in more than 5200 other sequenced bovine genomes. We speculate that the mutation occurred either as a parental germline mutation or post-zygotically in the developing embryo. This study implicates an important role for *KCNG1* as a member of the potassium voltage-gated channel group in neurodegeneration. Providing the first possible *KCNG1*-related disease model, we have, therefore, identified a new potential candidate for related conditions both in animals and in humans. This study illustrates the enormous potential of phenotypically well-studied spontaneous mutants in domestic animals to provide new insights into the function of individual genes.

Keywords: cattle; channelopathy; skeletal muscle; neuromuscular disorder; paradoxical myotonia congenita; potassium voltage-gated channel; precision medicine; hydrosyringomyelia; craniofacial dysmorphism

Citation: Jacinto, J.G.P.; Häfliger, I.M.; Akyürek, E.E.; Sacchetto, R.; Benazzi, C.; Gentile, A.; Drögemüller, C. *KCNG1*-Related Syndromic Form of Congenital Neuromuscular Channelopathy in a Crossbred Calf. *Genes* **2021**, *12*, 1792. https://doi.org/10.3390/genes12111792

Academic Editor: Emilia Bagnicka

Received: 18 October 2021
Accepted: 10 November 2021
Published: 12 November 2021

1. Introduction

Inherited channelopathies represent a clinically and heritably heterogeneous group of genetic disorders that result from a ion channel dysfunction of all cellular plasma membranes and/or of cell organelles [1]. They usually follow a dominant inheritance [1].

Neuromuscular channelopathies can cause different diseases affecting the brain, spinal cord, peripheral nerve, and/or muscle [2]. In particular, those that lead to primary skeletal muscle diseases, the so-called skeletal muscle channelopathies, exhibit a clinical spectrum ranging from flaccid paralysis to myotonia, the latter defined as delayed relaxation of a muscle that has been voluntarily or reflexively contracted [3].

In human medicine, skeletal muscle channelopathies are associated with pathogenic variants in genes coding for ion channels that influence muscle excitability [1,4]. They

are subdivided into two types, periodic paralysis (PPs) and non-dystrophic myotonias (NDMs) [5].

Various forms of PPs are due to abnormal depolarization that inactivates sodium channels, causing decreased muscle excitability of the muscle membrane, often accompanied by changes in extracellular potassium. The main finding is the susceptibility to episodes of focal or generalized weakness and paralysis [1]. This group of channelopathies includes three conditions: (a) hypokalemic periodic paralysis (HypoPP); (b) hyperkalemic periodic paralysis (HyperPP); and (c) Andersen-Tawil syndrome. HypoPP is the most common PP in humans, characterized by episodes of flaccid muscle weakness associated with low serum potassium levels, lasting hours to days [6]. The attacks of flaccid paralysis typically occur after waking during the night or early morning. It is due to a dysfunction of calcium channels associated with heterozygous variants in *CACNA1S* [7] or to a dysfunction of sodium channels associated with heterozygous variants in *SCN4A* [8]. HyperPP is characterized by episodes of flaccid muscle weakness associated with elevated serum potassium levels and occasional myotonia lasting minutes to hours [2]. It is due to dysfunction of sodium channels also associated with mutations in *SCN4A* [9]. The Andersen–Tawil syndrome is characterized by the following clinical triad: periodic paralysis, cardiac manifestations, and abnormal physical features. It is due to dysfunction of potassium channels associated with variants in *KCNJ2* [2,10]. In veterinary medicine, a *SCN4A*-related equine form of HyperPP has been reported (OMIA 000785-9796) [11].

Various forms of NDMs are due to defects in the muscle fiber repolarization, resulting in muscle hyperexcitability and myotonic discharges [10]. The main findings are muscular stiffness, in the absence of severe fixed weakness or muscle wasting, and muscle hypertrophy [4]. NDMs encompasses three different disorders, defined as following: (a) myotonia congenita (MC); (b) paramyotonia congenita (PMC); and (c) Na channel myotonias [6]. MC is the most common skeletal muscle channelopathy in humans, characterized by stiffness especially during rapid movements after a period of rest (action myotonia), and improving with exercise ("warm-up phenomenon") [3]. It is due to a dysfunction of chloride channels associated with dominant (Thomsen's disease) or recessively (Becker's disease) inherited mutations in *CLCN1* [12]. PMC is characterized by stiffness that, unlike MC, worsens with sustained exercise (exercise-induced or paradoxical myotonia) [6]. The symptoms last for seconds to minutes following the exercise. It is due to a dysfunction of sodium channels [13] also associated with heterozygous pathogenic variants in *SCN4A* [14]. There are several subtypes of sodium channel myotonia, such as acetazolamide-responsive myotonia, myotonia that develops approximately 10–20 min after exercise (*myotonia fluctans)* [15], and severe persistent myotonia associated with unique electromyographic pattern (*myotonia permanens)* [16]. Common to these subtypes is exacerbation by K (potassium-aggravated myotonias). Similarly to the PMC, they are all also associated with heterozygous pathogenic variants in *SCN4A* [14].

In veterinary medicine, forms of MC have been reported in horses (OMIA 000698-9796) [17], dogs (OMIA 000698-9615) [18], cats (OMIA 000698-9685) [19], sheep (OMIA 000698-9940) [20], and goats (OMIA 000698-9925) [21], associated with pathogenic variants in the orthologue *CLCN1* genes.

Altogether, it can be said that many clinicopathological similarities to *CLCN1*- and *SCN4A*-related human genetic diseases can be evidenced in veterinary pathology, highlighting the usefulness of translational research in the field of the congenital neuromuscular channelopathies. To our knowledge, no neuromuscular channelopathies have been reported in cattle. Therefore, with the present study we intended to characterize the clinical and pathological phenotype of a crossbred calf affected by congenital paradoxical myotonia, craniofacial dysmorphism, and myelodysplasia, and to find a possible genetic explanation after whole-genome sequencing (WGS).

2. Materials and Methods

2.1. Clinical and Pathological Investigation

A five-day-old male Belgian Blue x Holstein crossbred calf, weighting 47 kg, was admitted to the University of Bologna due to difficulty on quadrupedal stance and locomotion due to generalized muscle stiffness present since birth. The affected calf was clinically examined and a complete blood count (CBC), serum biochemical analysis, and venous blood gas analysis were obtained. Blood gas and serum biochemical analysis were performed at rest and after stimulation. Stimulation was the term used when the calf was in quadrupedal stance.

Nineteen days after hospitalization the calf showed a worsening of the general condition related to neuromuscular disease and was euthanized for welfare reasons. The calf was subsequently submitted for necropsy and histologic examination. Semimembranosus muscle was fixed in buffered neutral paraformaldehyde at 4 °C, washed in phosphate-buffered saline and de-hydrated through a graded series of ethanol. Samples embedded in paraffin were cut at 5 µm and stained with hematoxylin and eosin (H&E), or Azan–Mallory method, specific for detection of collagen fibers. Muscle sections were scanned with a semiautomatic microscope equipped (D-Sight v2, Menarini Diagnostics, Florence, Italy) with a computer. The average percentage of pathological muscle fibers was determined as the ratio of muscle fibers that lost their main shape and/or took a round shape to the total muscle fibers in the region. The spinal cord was fixed in 10% buffered formalin, embedded in paraffin, cut at 4 µm, and stained with hematoxylin and eosin (H&E), Periodic acid-Schiff (PAS), and Luxol-Fast-Blue for histological evaluation.

2.2. DNA Extraction, Whole-Genome Sequencing and Variant Calling

Genomic DNA was isolated from EDTA blood of the affected calf using a Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland). Using genomic DNA from the affected calf, an individual PCR-free fragment library with approximately 400 bp inserts was created and sequenced on a NovaSeq6000 for 150 bp paired-end reads (Illumina, San Diego, CA, USA). The sequenced reads were aligned to the ARS-UCD1.2 reference genome, resulting in an average coverage of approximately 17.4× [22], and single-nucleotide variants (SNVs) and small indel variants were called. The applied software and steps to process fastq files into binary alignment map (BAM) and genomic variant call format (GVCF) files were in accordance with the processing guidelines of the 1000 Bull Genomes Project (run 7) [23], with the exception of trimming, which was performed with fastp [24]. Further processing of the genomic data was performed according to Häfliger et al. 2020 [25]. The effects of the above variants were functionally evaluated with snpeff v4.3 [26], using the NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/; accessed on 17 July 2021). This resulted in the final VCF file, containing individual variants and their functional annotations. To find private variants, we compared the genotypes of the case with 691 cattle genomes of different breeds sequenced as part of the ongoing Swiss Comparative Bovine Resequencing project. All of its data are available (Table S1; https://www.ebi.ac.uk/ena/browser/view/PRJEB18113 accessed on 17 July 2021) in the European Nucleotide Archive (SAMEA7690196 is the sample accession number of the affected calf). Integrative Genomics Viewer (IGV) [27] software version 2.0 was used for visual evaluation of genome regions containing potential candidate genes.

2.3. Validation and Selection of Potential Canidate Variants

2.3.1. Occurrence of Variants in a Global Control Cohort

The comprehensive variant catalogue of run 9 of the 1000 Bull Genomes Project was available to investigate the allele distribution of variants within a global control cohort (www.1000bullgenomes.com; accessed on 17 July 2021) [23]. The whole data set includes 5116 cattle genomes including 576 from the Swiss Comparative Bovine Resequencing project, from a variety of breeds (>130 breeds indicated). Within the dataset, there are 9

purebred Belgian Blue and 1209 purebred Holstein cattle, allowing for the exclusion of common variants.

2.3.2. In Silico Assessment of the Molecular Consequences of Amino Acid Exchanges

Mutpred2 [28], PROVEAN [29] and PredictSNP1 [30] were used to predict the biological consequences of the detected missense variant. For cross-species sequence alignments, the following NCBI protein accessions were considered: NP_001192648.1 (*Bos taurus*), NP_002228.2 (*Homo sapiens*), XP_001168521.2 (*Pan troglodytes*), XP_543053.2 (*Canis lupus*), NP_001074603.1 (*Mus musculus*), NP_001100015.1 (*Rattus norvegicus*), XP_004947317.1 (*Gallus gallus*), and NP_001103880.1 (*Danio rerio*), NP_001096675.1 (*Xenopus tropicalis*).

2.4. Sequence Accessions

All references to the bovine *KCNG1* gene correspond to the NCBI accessions NC_037340.1 (chromosome 13, ARS-UCD1.2), NM_001205719.1 (*KCNG1* mRNA), and NP_001192648.1 (KCNG1 protein). For the protein structure of KCNG1 the Uniprot database (https://www.uniprot.org/; accessed on 17 July 2021) with accession number Q9UIX4 was used.

3. Results

3.1. Clinical Phenotype

On clinical examination at the time of admission, the calf was bright and alert but with generalized muscle stiffness that prevented it from spontaneously assuming and maintaining the quadrupedal stance.

At rest, the animal preferred the sternal recumbency, with the forelimbs folded under its chest while the hindlimbs were rigid and hyperextended (Figure 1a). It was not possible to flex the hindlimbs due to the muscle stiffness. If stimulated to stand the muscle stiffness increased inducing a rigid posture accompanied by spastic contractions that prevented him to acquire a definitive quadrupedal stance. On the contrary, if gently passively positioned, the calf was able to acquire and maintain the quadrupedal stance. In standing, the hind limbs were contracted and hyperextended, especially the right hindlimb that showed caudal stretching (Video S1). Additionally, the back was slightly arched, and the tail head elevated (Figure 1b). The thoracic girdle was also involved but less severely. On hooping and hoof replacement, the calf was unable to re-acquire the physiological position of the limbs/hoof. Unless supported, the calf was unable to walk or maintain the quadrupedal stance for long time. In fact, uncontrolled hypertonic postural reactions and muscular contractions resulted in loss of stance, with a fall in lateral recumbency. If not further stimulated and stressed, the stiffness slowly tended to decrease, enabling the calf to acquire the sternal recumbency. However, the hypertonia never disappeared completely. The cutaneous trunci reflex was increased in intensity as well as the withdrawal reflex of the forelimbs while in the hindlimbs the latter was absent. No abnormalities in the cranial nerves' reflexes, threat response, and pain perception were noticed.

The calf showed a slight flattening of the splanchnocranium with deviation to the right side. It displayed carpal and tarsal skin lesions due to permanent recumbency. Moreover, the animal presented diarrhea.

CBC revealed moderate leukocytosis (23,650/mm^3) with neutrophilia (14,280/mm^3) and monocytosis (2400/mm^3). Serum biochemical profile after stimulation showed increase in: creatinine kinase, lactate dehydrogenase (LDH), L-lactate, potassium (K^+), and calcium (Ca^{2+}) (Table 1).

Based on the clinical findings, the calf was suspected of suffering from a form of paradoxical myotonia congenita and/or from a spinal cord lesion associated with craniofacial dysmorphism.

Figure 1. Crossbred calf with a congenital neuromuscular disorder characterized by paradoxical myotonia congenita and myelodisplasia. (**a**) Sternal recumbency at rest. Note that the calf sustains the forelimbs folded underneath its chest while the hindlimbs are hyperextended. (**b**) Quadrupedal stance after passive positioning. Note that the pelvic girdle appears to be more affected with a marked hyperextension of the hindlimbs, the back is slightly arched, and the tail head is elevated. The thoracic girdle was affected with the forelimbs' hoofs resting on tip.

Table 1. Results of altered parameters of serum biochemical profiles of the affected calf at rest and after stimulation.

Parameter	At Rest	After Stimulation	Unit of Measure
Creatinine kinase (CK)	147	329	IU/L
Lactate dehydrogenase (LDH)	1975	2556	IU/L
L-lactate	1	3.8	mmol/L
Potassium (K+)	3.6	4.3	mmol/L
Calcium (Ca2+)	0.99	1.14	mmol/L

3.2. Pathological Phenotype

At gross pathology, the examination of central nervous system revealed narrowing of the spinal cord (myelodysplasia) between lumbar spinal nerves IV and VI associated with hydrosyringomyelia between lumbar spinal nerves III and V with the larger cavity at the level of lumbar spinal nerve IV (Figure 2a,b). Macroscopically, the muscles were normal.

Microscopically, at lumbar spinal nerves III to V there were two cavities with only the larger partially lined by ependymal cells (hydrosyringomielia) (Figure 2c). Vasogenic and intramyelinic edemas were present in the areas around the cavities, characterized as multiple fluid-filled clear extracellular spaces in the gray matter. Around the capillaries near the smaller channel, there was neutrophilic and lymphoid inflammatory cells. The astrocytes in the white matter showed foci of chromatin margination.

Routinely morphological (hematoxylin-eosin) analysis was used for histopathological evaluation on semimembranosus muscle biopsy sections. Muscle parenchyma showed normal fibers distribution. (Figure 3a). Nevertheless, some fibers appeared round shaped (Figure 3c) and most of them exhibited an enlarged cross-sectional area (Figure 3d). It was determined that the average percentage of pathological muscle fibers was 5.3%. A possible presence of fibrosis was determined by Azan–Mallory staining (Figure 3b). No signs of fibrosis were found in the sections obtained from the tissue. Infiltrated inflammatory mononuclear cells were not revealed by histological analysis of the muscle.

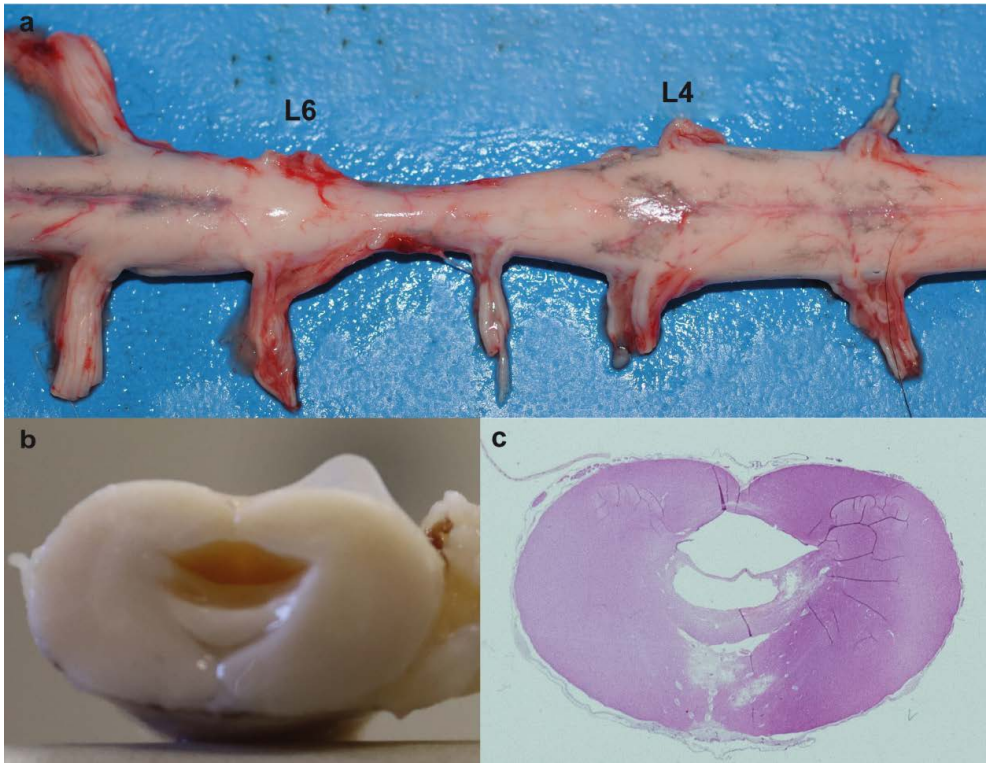


Figure 2. Myelodysplasia associated with hydrosyringomyelia in the affected calf. (**a**) Note the narrowing of the spinal cord between lumbar spinal nerves IV (L4) and VI (L6) (myelodysplasia). (**b**) Transversal section of the spinal cord between lumbar spinal nerve V (L5) and VI (L6). Note the cavity formed within the spinal cord. (**c**) Histological section of (**b**). Note that there are two cavities with only the larger partially lined by ependymal cells (hydrosyringomielia). hematoxylin and eosin (H&E) staining.

3.3. Genetic Analysis

Assuming spontaneous mutation as the cause of this congenital neuromuscular condition, the WGS data were filtered for heterozygous coding variants that were present in the calf and were absent in the 691 available cattle genomes of different breeds. Thereby, 151 variants with a predicted high or moderate impact were identified (Table 2). In a second step, these variants were analyzed for their occurrence in a global cohort of 4540 genomes from a variety of breeds. This revealed 27 remaining protein-changing variants that are exclusively heterozygous in the affected calf and absent in all controls. These 27 variants were then visually inspected using the IGV software (Broad institute, Cambridge, MA, USA), which confirmed 25 as true variants (Tables 2 and S2).

Table 2. Results of whole-genome sequencing variant filtering of the calf affected by paramyotonia congenita and myelodysplasia.

Filtering Step	Homozygous Variants	Heterozygous Variants
All variants	2,562,043	5,168,233
Private variants	3580	21,104
Protein-changing private variants using 691 cattle genome controls	12	115
Remaining protein-changing private variants using a global control cohort of 4540 cattle genomes and subsequent IGV inspection	0	25

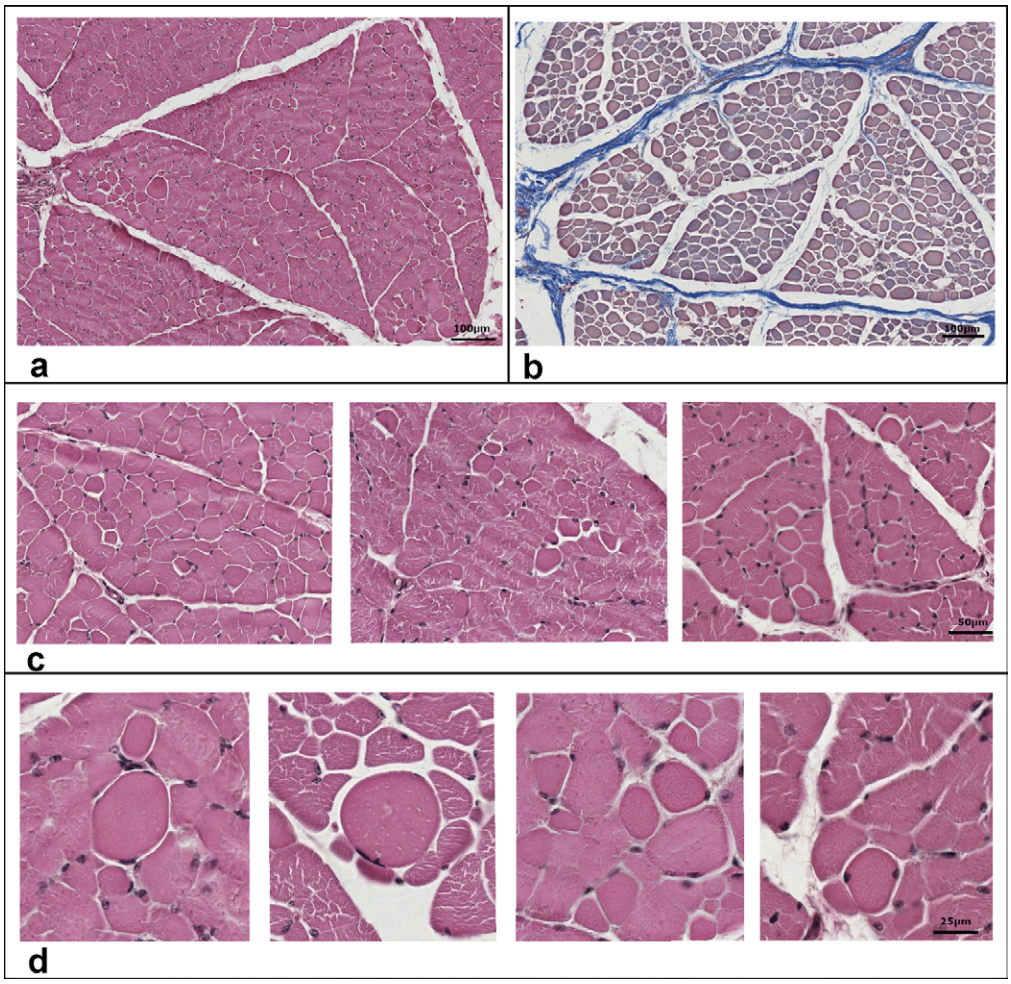


Figure 3. Histological features of semimembranosus muscle of the studied case. Transversal sections from muscle biopsies were stained with hematoxylin and eosin (**a**,**c**,**d**) or with Azan–Mallory method (**b**) to identify collagen fibers. Enlarged round shaped fibers are highlighted in panels (**c**,**d**). The percentage value of round shaped fibers (5.3%) was determined as the ratio of muscle fibers that lost their main shape and/or took a round shape to the total muscle fibers in the region. Scale bars correspond to 100 m in panels (**a**,**b**), and 50 and 25 m in panels c and d, respectively.

Among these 25 remaining private variants, one single variant affects an interesting functional candidate gene for the studied phenotype (Figure 4a). This heterozygous variant at chr13:78918850C>A is a missense variant in exon 2 of the *potassium voltage-gated channel modifier subfamily G member 1* (*KCNG1*) gene (NM_001205719.1: c.1248G>T; Figure 4b). It exchanges the encoded amino acid of *KCNG1* at position 416 (NP_001192648.1: p.Trp416Cys), located in the pore helix of the ion transport domain (Figure 4c). Furthermore, the tryptophan-to-cysteine substitution affects a highly conserved residue (4d), and was predicted to be deleterious by three different tools (Mutpred2 score 0.951; PROVEAN score -12.212; PredictSNP1 score 0.869) and to alter the transmembrane protein and ordered interface. Unfortunately, biological samples of the dam and sire that were slaughtered in the meantime, were not available. Analysis of the other 24 identified variants, taking into account the known function of the gene, the reported association with Mendelian diseases, and/or the in silico assessment of the molecular consequences of the variants in the protein, did not reveal any other plausible cause for the observed phenotype (Table S2).

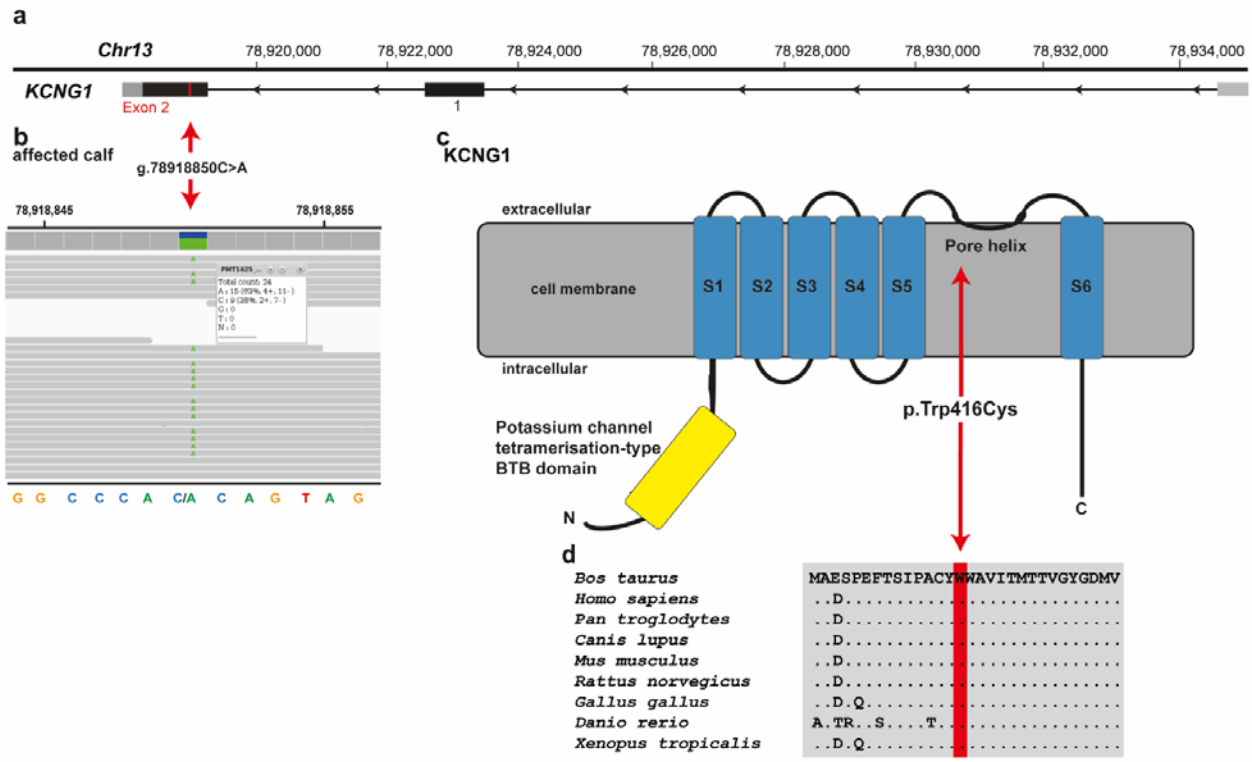


Figure 4. *KCNG1* missense variant in a crossbred calf with paradoxical myotonia congenita and myelodysplasia. (**a**) Structure of *KCNG1* showing the exon 2 variant located on chromosome 13. (**b**) IGV screenshot presenting the Chr13: 78918850C>A variant in the affected calf. (**c**) Schematic representation of *KCNG1* protein and its functional domains with the position of the identified pathogenic variant (red arrow). The six transmembrane domains are shown in blue (S1-S6). (**d**) Cross-species sequence comparison of the ion transport domain of the KCNG1protein with the region around the p.Trp416Cys variant shows complete evolutionary conservation.

Assuming recessive inheritance, filtering of WGS data for homozygous coding variants present in the calf and missing in the 691 control genomes of different breeds identified 12 variants with likely moderate impact. These 12 variants were further investigated for their occurrence in a diverse cohort of additional 4540 bovine genomes, which revealed the absence of variants only present homozygous in the affected calf (Table 2).

4. Discussion

This study aimed to investigate the clinicopathological phenotype and the underlying genetic cause in a crossbred calf displaying paradoxical myotonia congenita, craniofacial dysmorphism and myelodysplasia. The phenotype of this syndromic form of congenital neuromuscular disorder displays striking similarities with *SCN4A*-related forms of human PMC because it presents with episodes of exercise-induced generalized myotonic muscle stiffness. Clinical diagnosis of PMC in humans is based on consistent history, and typical clinical and electromyographic findings [6].

In PMC patients, changes in muscle fiber diameters and internal nuclei are also among the nonspecific histologic features suggestive of mild myopathic changes [31]. The muscle profile in this study revealed that 5.3% of muscle fibers lost their main shape and/or took a round shape to the total muscle fibers in the region, findings that are consistent with the human PMC features. However, the lack of an electromyography prevented a definite categorization of this paradoxical myotonia as PMC.

Moreover, some further phenotypical differences from the typical form of human PMC were noticed, such as: (1) increase in serum K^+ after stimulation; (2) inability to relax muscle immediately after the stimulation; (3) craniofacial dysmorphism; and (4) permanent extension of the hindlimbs, the latter explained by the retrieved myelodysplasia associated

with hydrosyringomyelia in the lumbosacral intumescence region. Taken together, the muscle stiffness episodes and the findings of points (1) and (2) more resembled *CACNA1S*-related forms of HyperPP. Human patients with this genetic disease show an increase in serum potassium during an attack [4]. Moreover, approximately half of the patients display muscle stiffness arising from myotonia or paramyotonia [32]. On the other side, considering the clinical muscle findings and point (3)—craniofacial dysmorphism—the observed phenotype shows similarities to the human *KCNJ2*-related Andersen–Tawil syndrome.

Hence, to the best of our understanding, our patient showed a previously unreported combination of paradoxical myotonia congenita, hyperkalemia during episodes, craniofacial dysmorphism, and myelodysplasia associated with hydrosyringomyelia, representing a novel clinicopathological presentation.

In humans, genetic confirmation of known pathogenic variants in *SCN4A*-related PMC and HyperPP, and *KCNJ2*-related Andersen–Tawil syndrome-related, is included in the diagnosis of these disorders [6]. In the studied calf, no candidate causal variant in *CACNA1S*, *SCN4A*, or *KCNJ2* were found by genome re-sequencing. Likewise, in human medicine, patients with phenotypical characteristics of PMC, HyperPP, and Andersen–Tawil syndrome were found not to present pathogenic variants in the *SCN4A* or in *KCNJ2*, suggesting further genetic heterogeneity [33]. Therefore, we evaluated the possible genetic cause for this novel congenital phenotype systematically, assuming both recessively inherited and de novo mutations. Our results from the analysis of WGS data showed that there was not a single homozygous protein-changing variant present in the affected calf, ruling out a possible recessive inheritance as the most likely cause. Furthermore, as our case was an offspring of a crossbred mating it seems to be highly unlikely that a monogenic recessive variant was causal. Especially since no bovine simple genetic disease is known that segregates in such diverse dairy and beef breeds as Holstein and Belgian-Blue, respectively. Therefore, the more plausible explanation would be to search for allelic heterogeneity, meaning two different (breed-specific) coding variants affecting the same gene. Our results from the WGS identified 25 heterozygous private protein-changing variants present in the genome of the affected calf which were absent in a cohort of more than 5200 cattle genomes. Considering the known function of the affected gene, the rarity of the variant, and the outcome of the in silico effect prediction, the identified heterozygous *KCNG1* missense variant was assumed to represent the most likely genetic cause for the observed phenotype. We could only speculate that the mutation either occurred post-zygotically in the developing embryo or it represents a germline mutation in the dam or sire. To confirm that the identified mutation in the *KCNG1* gene occurred indeed de novo, genotyping of the parents would be needed. Unfortunately, no genetic material of both parents was available. To the best of our knowledge, no pathogenic variant in *KCNG1* gene has been reported in animals and humans. Therefore, this study represents the first example of a *KCNG1*-related neuromuscular disorder as a conserved residue in the pore helix of the ion transport domain of the potassium voltage-gated channel subfamily G member 1 protein is altered.

Voltage-gated potassium channels represent a family of transmembrane proteins that are highly expressed in the central nervous system of the mammalian species, playing a major role in the control of neuronal excitability [34]. Additionally, they regulate a variety of electrophysiological properties, including the interspike membrane, the wave-form of the action potential and the firing frequency [35]. In particular, *KCNG1* is a potassium channel subunit that cannot form functional channels by itself [36]. However, it forms functional channels with the *KCNB1* (one of the most important voltage-gated potassium channels) and modulates the delayer rectifier voltage-gated potassium channel activation and deactivation rated of this protein [34,37]. Specifically, *KCNG1* in co-expression with *KCNB1* results in potassium channels that have slower kinetics of deactivation, due to a negative shift of the steady-state activation curve, and marked slowing of deactivation tail currents [34,37]. The performed in silico evaluations of the identified p.Trp416Cys mutation predicted that this mutation altered the transmembrane protein and ordered

interface. Therefore, we hypothesize that our mutant protein leads to a malfunction of the encoded channels allowing an additional release of intracellular potassium from the skeletal muscle. Subsequently, this change in ion transport impairs the ability of the muscle to contract, leading to the observed stiffness.

Moreover, in human medicine, there are several potassium channelopathies whose presentations are suggestive of developmental disorders, with findings including intellectual disability, craniofacial dysmorphism or other physical abnormalities [38]. Physiological functions of *KCNG1* are largely unknown, by similarity, therefore, it is plausible that the identified mutation in our study is also the underlying cause for the craniofacial dysmorphism and myelodysplasia associated with hydrosyringomyelia.

5. Conclusions

We have uncovered a novel phenotype of a most likely dominantly inherited neuromuscular channelopathy in cattle related to a potentially pathogenic variant in the bovine *KCNG1* gene. Targeted expression of affected potassium channels in transfected cell lines, in combination with recently developed gene editing tools such as CRISPR/Cas9 to mimic the consequences of the exchanged residue, may be suitable to functionally prove our claims in the future. Nevertheless, we propose here the first *KCNG1* mutation present with a disorder in a mammalian species. Our study highlights that the genetics of inherited disorders in well-phenotyped large animals, such as cattle, is a valuable model system for studying fundamental aspects of gene function.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/genes12111792/s1, Table S1: EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 691 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available; Table S2: List of the remaining variants after the comparison to the global control cohort of 5116 genomes of other breeds (1000 Bull Genomes Project run 9; www.1000bullgenomes.com; acceded on 17 July 2021) and after IGV visual inspection, revealing 25 protein-changing variants with a predicted moderate or high impact only present in the affected calf; Video S1: After passively positioned, the calf was able to acquire and maintain the quadrupedal stance. In standing the hind limbs were contracted and hyperextended, and tail head elevated.

Author Contributions: Conceptualization, C.D. and A.G.; methodology, I.M.H., E.E.A., R.S., C.B., A.G. and C.D.; validation, J.G.P.J.; formal analysis, J.G.P.J., I.M.H., E.E.A., R.S., C.B., A.G. and C.D.; investigation, J.G.P.J. and A.G.; resources, C.D.; data curation, I.M.H.; writing—original draft preparation, J.G.P.J.; writing—review and editing, J.G.P.J., I.M.H., E.E.A., R.S., C.B., A.G. and C.D.; visualization, R.S., A.G. and J.G.P.J.; supervision, C.D. and A.G.; project administration, C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by the Swiss National Science Foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but part of a clinical and pathological veterinary diagnostic case.

Informed Consent Statement: Not applicable.

Data Availability Statement: The whole-genome data of the affected calf (sample ID PMT1625) is freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA7690196.

Acknowledgments: We acknowledge the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kim, J.-B. Channelopathies. *Korean J. Pediatr.* **2014**, *57*, 1–18. [CrossRef] [PubMed]
2. Spillane, J.; Kullmann, D.M.; Hanna, M.G. Genetic neurological channelopathies: Molecular genetics and clinical phenotypes. *J. Neurol. Neurosurg. Psychiatry* **2016**, *87*, 37–48. [CrossRef] [PubMed]

3. Colding-Jørgensen, E. Phenotypic variability in myotonia congenita. *Muscle Nerve* **2005**, *32*, 19–34. [CrossRef]
4. Phillips, L.; Trivedi, J.R. Skeletal Muscle Channelopathies. *Neurother. J. Am. Soc. Exp. Neurother.* **2018**, *15*, 954–965. [CrossRef]
5. Montagnese, F.; Schoser, B. Dystrophic and non-dystrophic myotonias. *Fortschr. Neurol. Psychiatr.* **2018**, *86*, 575–583. [CrossRef] [PubMed]
6. Stunnenberg, B.C.; LoRusso, S.; Arnold, W.D.; Barohn, R.J.; Cannon, S.C.; Fontaine, B.; Griggs, R.C.; Hanna, M.G.; Matthews, E.; Meola, G.; et al. Guidelines on clinical presentation and management of nondystrophic myotonias. *Muscle Nerve* **2020**, *62*, 430–444. [CrossRef] [PubMed]
7. Fontaine, B.; Vale-Santos, J.; Jurkat-Rott, K.; Reboul, J.; Plassart, E.; Rime, C.S.; Elbaz, A.; Heine, R.; Guimarães, J.; Weissenbach, J. Mapping of the hypokalaemic periodic paralysis (HypoPP) locus to chromosome 1q31-32 in three European families. *Nat. Genet.* **1994**, *6*, 267–272. [CrossRef] [PubMed]
8. Bulman, D.E.; Scoggan, K.A.; van Oene, M.D.; Nicolle, M.W.; Hahn, A.F.; Tollar, L.L.; Ebers, G.C. A novel sodium channel mutation in a family with hypokalemic periodic paralysis. *Neurology* **1999**, *53*, 1932–1936. [CrossRef]
9. Ptacek, L.J.; Trimmer, J.S.; Agnew, W.S.; Roberts, J.W.; Petajan, J.H.; Leppert, M. Paramyotonia congenita and hyperkalemic periodic paralysis map to the same sodium-channel gene locus. *Am. J. Hum. Genet.* **1991**, *49*, 851–854. [PubMed]
10. Raja Rayan, D.L.; Hanna, M.G. Skeletal muscle channelopathies: Nondystrophic myotonias and periodic paralysis. *Curr. Opin. Neurol.* **2010**, *23*, 466–476. [CrossRef] [PubMed]
11. Koch, M.C.; Steinmeyer, K.; Lorenz, C.; Ricker, K.; Wolf, F.; Otto, M.; Zoll, B.; Lehmann-Horn, F.; Grzeschik, K.H.; Jentsch, T.J. The skeletal muscle chloride channel in dominant and recessive human myotonia. *Science* **1992**, *257*, 797–800. [CrossRef] [PubMed]
12. Cannon, S.C. Channelopathies of skeletal muscle excitability. *Compr. Physiol.* **2015**, *5*, 761–790. [CrossRef] [PubMed]
13. Ptácek, L.J.; George, A.L.J.; Barchi, R.L.; Griggs, R.C.; Riggs, J.E.; Robertson, M.; Leppert, M.F. Mutations in an S4 segment of the adult skeletal muscle sodium channel cause paramyotonia congenita. *Neuron* **1992**, *8*, 891–897. [CrossRef]
14. Ricker, K.; Moxley, R.T., 3rd; Heine, R.; Lehmann-Horn, F. *Myotonia fluctuans*. A third type of muscle sodium channel disease. *Arch. Neurol.* **1994**, *51*, 1095–1102. [CrossRef] [PubMed]
15. Colding-Jørgensen, E.; Duno, M.; Vissing, J. Autosomal dominant monosymptomatic myotonia permanens. *Neurology* **2006**, *67*, 153–155. [CrossRef]
16. Wijnberg, I.D.; Owczarek-Lipska, M.; Sacchetto, R.; Mascarello, F.; Pascoli, F.; Grünberg, W.; van der Kolk, J.H.; Drögemüller, C. A missense mutation in the skeletal muscle chloride channel 1 (CLCN1) as candidate causal mutation for congenital myotonia in a New Forest pony. *Neuromuscul. Disord.* **2012**, *22*, 361–367. [CrossRef]
17. Rhodes, T.H.; Vite, C.H.; Giger, U.; Patterson, D.F.; Fahlke, C.; George, A.L.J. A missense mutation in canine C1C-1 causes recessive myotonia congenita in the dog. *FEBS Lett.* **1999**, *456*, 54–58. [CrossRef]
18. Gandolfi, B.; Daniel, R.J.; O'Brien, D.P.; Guo, L.T.; Youngs, M.D.; Leach, S.B.; Jones, B.R.; Shelton, G.D.; Lyons, L.A. A novel mutation in CLCN1 associated with feline myotonia congenita. *PLoS ONE* **2014**, *9*, e109926. [CrossRef]
19. Monteagudo, L.V.; Tejedor, M.T.; Ramos, J.J.; Lacasta, D.; Ferrer, L.M. Ovine congenital myotonia associated with a mutation in the muscle chloride channel gene. *Vet. J.* **2015**, *204*, 128–129. [CrossRef] [PubMed]
20. Beck, C.L.; Fahlke, C.; George, A.L.J. Molecular basis for decreased muscle chloride conductance in the myotonic goat. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11248–11252. [CrossRef] [PubMed]
21. Rudolph, J.A.; Spier, S.J.; Byrns, G.; Rojas, C.V.; Bernoco, D.; Hoffman, E.P. Periodic paralysis in quarter horses: A sodium channel mutation disseminated by selective breeding. *Nat. Genet.* **1992**, *2*, 144–147. [CrossRef] [PubMed]
22. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, giaa021. [CrossRef] [PubMed]
23. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef] [PubMed]
24. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef]
25. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef]
26. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef]
27. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant review with the integrative genomics viewer. *Cancer Res.* **2017**, *77*, e31–e34. [CrossRef] [PubMed]
28. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.; Lin, G.N.; Nam, H.-J.; Mort, M.; Cooper, D.; Sebat, J.; Iakoucheva, L.; et al. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv* **2017**. [CrossRef]
29. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]
30. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput. Biol.* **2014**, *10*, e1003440. [CrossRef] [PubMed]

31. Izumi, Y.; Fukuuchi, Y.; Koto, A.; Nakajima, S. Paramyotonia congenita without cold paralysis: A case report. *Keio J. Med.* **1994**, *43*, 94–97. [CrossRef] [PubMed]
32. Statland, J.M.; Fontaine, B.; Hanna, M.G.; Johnson, N.E.; Kissel, J.T.; Sansone, V.A.; Shieh, P.B.; Tawil, R.N.; Trivedi, J.; Cannon, S.C.; et al. Review of the Diagnosis and Treatment of Periodic Paralysis. *Muscle Nerve* **2018**, *57*, 522–530. [CrossRef] [PubMed]
33. Miller, T.M.; Dias da Silva, M.R.; Miller, H.A.; Kwiecinski, H.; Mendell, J.R.; Tawil, R.; McManis, P.; Griggs, R.C.; Angelini, C.; Servidei, S.; et al. Correlating phenotype and genotype in the periodic paralyses. *Neurology* **2004**, *63*, 1647–1655. [CrossRef] [PubMed]
34. Schnitzler, M.M.; Rinné, S.; Skrobek, L.; Renigunta, V.; Schlichthörl, G.; Derst, C.; Gudermann, T.; Daut, J.; Preisig-Müller, R. Mutation of histidine 105 in the T1 domain of the potassium channel Kv2.1 disrupts heteromerization with Kv6.3 and Kv6.4. *J. Biol. Chem.* **2009**, *284*, 4695–4704. [CrossRef]
35. Hille, B. *Ion. Channels of Excitable Membranes*, 3rd ed.; Sinauer Associates Institute, Ed.; Sinauer: Sunderland, MA, USA, 2001.
36. Post, M.A.; Kirsch, G.E.; Brown, A.M. Kv2.1 and electrically silent Kv6.1 potassium channel subunits combine and express a novel current. *FEBS Lett.* **1996**, *399*, 177–182. [CrossRef]
37. Kramer, J.W.; Post, M.A.; Brown, A.M.; Kirsch, G.E. Modulation of potassium channel gating by coexpression of Kv2.1 with regulatory Kv5.1 or Kv6.1 α-subunits. *Am. J. Physiol. Physiol.* **1998**, *274*, C1501–C1510. [CrossRef] [PubMed]
38. Hamilton, M.J.; Suri, M. *Chapter Four—"Electrifying Dysmorphology": Potassium Channelopathies Causing Dysmorphic Syndromes*; Kumar, B.T.-A., Dionis, G., Eds.; Academic Press: New York, NY, USA, 2020; Volume 105, pp. 137–174. ISBN 0065-2660.

A *de novo* variant in the bovine *ADAMTSL4* gene in an Original Braunvieh calf with congenital cataract

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, visualizations, co-writing first draft, review and revisions
Displayed version:	published version
DOI:	10.1111/age.13178

Received: 27 January 2022 | Accepted: 3 February 2022

DOI: 10.1111/age.13178

SHORT COMMUNICATION

ANIMAL GENETICS WILEY

A *de novo* variant in the bovine *ADAMTSL4* gene in an Original Braunvieh calf with congenital cataract

Irene M. Häfliger[1] | **Sonja Wolf-Hofstetter**[1] | **Christina Casola**[2] | **Udo Hetzel**[3] | **Franz R. Seefried**[4] | **Cord Drögemüller**[1]

[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland

[2]Ophthalmology Section, Equine Department, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

[3]Institute of Veterinary Pathology, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

[4]Qualitas AG, Zug, Switzerland

Correspondence
Cord Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, Bern 3001, Switzerland.
Email: cord.droegemueller@vetsuisse.unibe.ch

Funding information
Swiss National Science Foundation, Grant/Award Number: 172911

Abstract

Inherited forms of cataract are a heterogeneous group of eye disorders known in livestock species. Clinicopathological analysis of a single case of impaired vision in a newborn Original Braunvieh calf revealed nuclear cataract. Whole-genome sequencing of the parent-offspring trio revealed a de novo mutation of *ADAMTSL4* in this case. The heterozygous p.Arg776His missense variant affects a conserved residue of the *ADAMTSL4* gene that encodes a secreted glycoprotein expressed in the lens throughout embryonic development. In humans, *ADAMTSL4* genetic variants cause recessively inherited forms of subluxation of the lens. Given that *ADAMTSL4* is a functional candidate gene for inherited disorders of the lens, we suggest that heterozygosity for the identified missense variant may have caused the congenital cataract in the affected calf. Cattle populations should be monitored for unexplained cataract cases, with subsequent DNA sequencing a hypothesized pathogenic effect of heterozygous *ADAMTSL4* variants could be confirmed.

KEYWORDS
Bos taurus, development, eye, lens, precision medicine, whole-genome sequencing

Congenital ocular conditions such as cataract causing impaired vision or complete blindness have been documented in cattle (Williams, 2010). It was established more than 100 years ago that cataracts in mammals can be environmental or hereditary (Detlefson & Yapp, 1920). For example, congenital ocular lesions including primarily cortical cataracts have been reported in calves exposed to bovine viral diarrhea virus in utero (Bistner et al., 1970; Siepker et al., 2021). However, often no definitive nutritional, infectious, or toxic cause for the congenital cataract is identified in affected herds (Krump et al., 2014; Osinchuk et al., 2017). Sporadic occurrence of inherited forms of cataract in certain breeds of cattle has been reported (OMIA 000168-9913). Detlefson and Yapp (1920) reported evidence of monogenic autosomal recessive inheritance in Holsteins and Gregory et al. (1943) reported equally convincing data in Jerseys. Nonetheless, so far, causative variants were found only for two rare breed-specific recessive forms of bovine cataract: the *NID1*-related nuclear cataract of Romagnolas (OMIA 001936-9913; Murgiano et al., 2014), and the *CPAMD8*-related Morgagnian cataract of Holsteins (OMIA 002111-9913; Hollmann et al., 2017). Such findings enable selection against these disorders within the affected populations. Furthermore, the cataract observed in Romagnolas was the first report of a naturally occurring mutation of *NID1* (Murgiano et al., 2014). This led to a non-syndromic form of cataract in

a mammalian species and added the affected gene to the list of candidate genes for inherited forms of nuclear cataract in humans (Murgiano et al., 2014). In human medicine, about 50% of cataracts are thought to be associated with genetic factors and causative variants for congenital or other early-onset forms of cataracts have been discovered in over 30 genes (Shiels & Hejtmancik, 2017). Congenital cataract has also been shown to be a heterogeneous group of diseases in cattle, as the reported *CPAMD8* nonsense variant is obviously not sufficient to explain the majority of Holsteins suffering from Morgagnian congenital cataract (Braun et al., 2019).

In the present study, we investigated a 1.5-month-old male Original Braunvieh calf with clinical signs of cataract showing bilateral opaque eye lenses (Figure 1). According to the owner, the cataract had been present since birth. At first, the calf had problems finding the teats to suckle the milk and initially ran into walls and objects, but after a few days, it was able to orientate itself. Clinical examination revealed no further abnormalities and the calf was referred to an ophthalmologist. The calf presented to the ophthalmology service with open eyes. Dazzle and pupillary light reflexes were normal, but menace response was absent in both eyes. Intraocular pressure was within normal limits (right eye 20 mmHg, left eye 19 mmHg). The lens showed in both eyes a diffuse dense nuclear opacity and a focal axial anterior cortical opacity in addition with multifocal anterior capsular punctual opacities. Fundic examination was mostly blocked by lens opacities, the visible parts in the periphery were normal. Further ophthalmic examination showed no more abnormalities and the calf was diagnosed with an immature cataract in both eyes. The parents of the affected calf were clinically normal. At the time of slaughtering, the calf was 4.5 months old and appeared in a good general health condition.

Macroscopically, the regularly shaped ocular bulbs demonstrated normal sized lenses with moderately diffuse clouding with clear demarcation of the capsules and regularly developed zonula fibers (Figure 2a,b,d). Histologically, the anterior half of lental protein appears normal, whereas from the equator onwards posteriorly, a severe lental nuclear cataract composed of eosinophilic partly fibrillary, partly amorphous lental proteinaceous material with multifocal small foci of mineralisation (calcification) is present (Figure 2c,e,f). The anterior lens capsule is well developed, the anterior epithelium and equator show a regular formation whereas the posterior lens capsule reveals a moderate fibrillation (Figure 2g, right side). No evidence for infectious agents could be observed in hematoxylin and eosin as well as special stains. All other intraocular structures appeared histologically normal. The bovine embryonal lental development occurs during the first trimester of gestation with formation of the lens placode to development of the lens vesicle between 3.3- and 14-mm embryonal crown rump length (Schnorr & Kressin, 2006). Lenses show continuous lifelong growth with strongest growth rate during embryonal and morphologically non-altered development and the 1st year of life (Levin et al., 2011). The lenses presented here are fully developed with an intact lental capsule, zonula fibers, and regularly arranged and morphologically normal equatorial cells. Thus the opacification of the lental protein, i.e. the formation of a nuclear and posterior subcapsular cataract, must have developed in a relative late stage of the lental development, morphologically widely comparable to the findings demonstrated in the *NID1*-related nuclear cataract of Romagnolas (Murgiano et al., 2014) and Morgagnian cataracts in Red Holstein Friesian cattle (Braun et al., 2019).

We prepared a PCR-free DNA library of the affected calf and its dam and collected short read pairs (2 × 150 bp) to obtain roughly 25× coverage on an Illumina NovaSeq 6000 instrument. The whole genome sequencing data of the sire was publicly available. Variants in the genome of the affected calf were called with respect to the reference genome assembly ARS-UCD1.2 as described previously (Häfliger et al., 2020). A hard filtering approach was applied by comparing the variants to a cohort of 5115 cattle genomes containing both parents. Given that the mode of inheritance was unknown and assuming that the causative variant is rare, we filtered for variants that were only present in the affected calf, either in heterozygous or homozygous state. All 5116 genomes are included in the variant catalogue of run 9 of the 1000 Bull Genomes Project including animals of more than 130 genetically diverse breeds (Hayes & Daetwyler, 2019). Assuming recessive

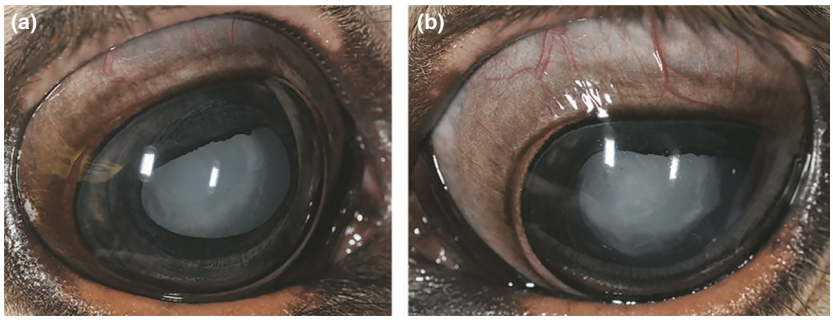


FIGURE 1 Original Braunvieh calf with congenital cataract. Opaque lenses of right (a) and left (b) eye are shown at the ophthalmological examination at 1.5 months

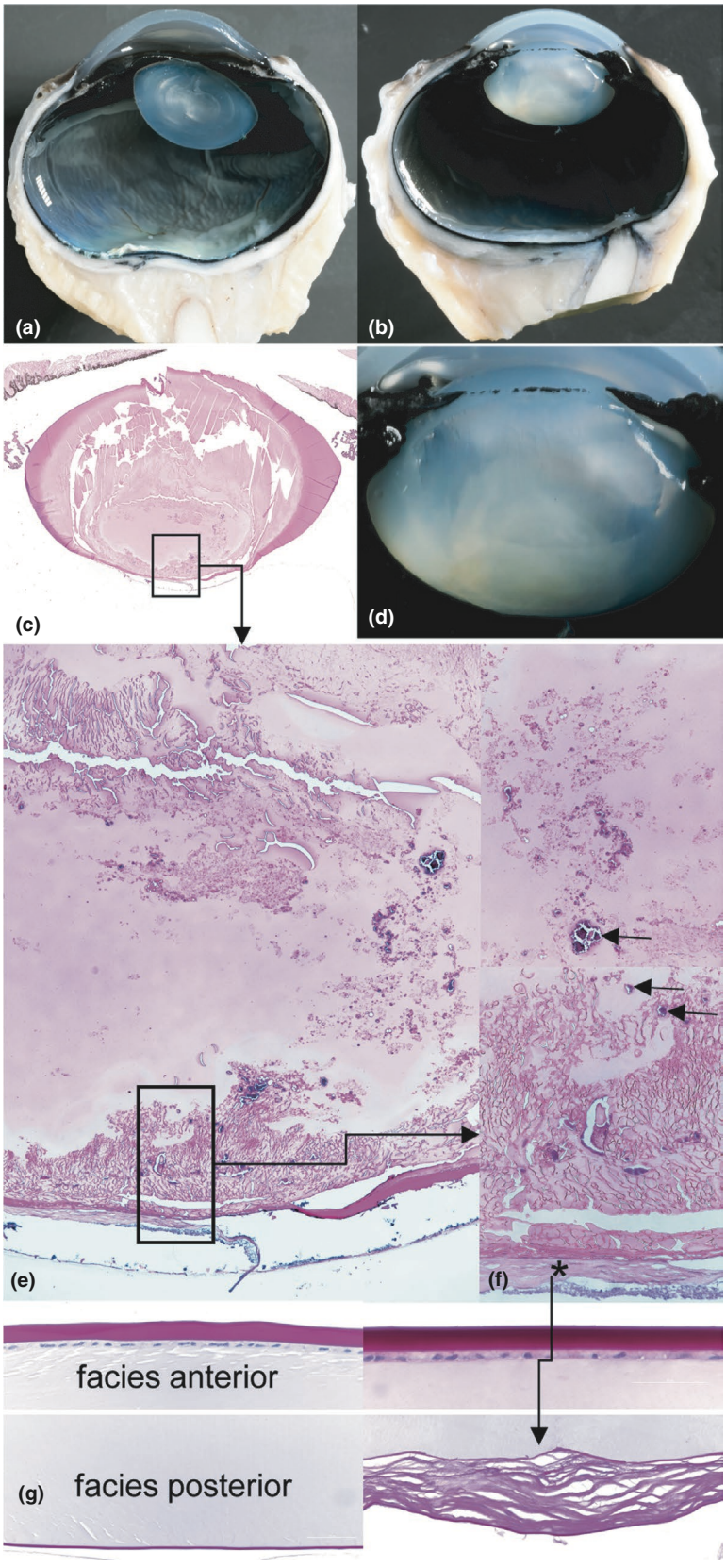


FIGURE 2 Features of nuclear cataract in an Original Braunvieh calf. Normal eye, sagittal section (a). Altered right eye with severe lental nuclear cataract (rectangle), sagittal section (b). Histological overview (c) of the cataractous lens depicted in (b). Higher magnification (d) of the lens depicted in (b). Facies posterior of the lens with cloudy appearance of lental protein and multiple foci of mineralization (e). Higher magnification of the posterior lens capsule, with perilental collagenous membrane (asterisk) and fibrillar lental protein with mineralization (arrows) (f). Comparison of the anterior and posterior lens capsule, left side: normal bovine lens, right: altered fibrillary posterior capsule (g)

inheritance, filtering for coding variants present only homozygous in the calf and heterozygous in the parental genomes identified no coding variant. Assuming a dominant mutation, filtering revealed a single private protein changing variant, a missense variant in the *ADAMTSL4* gene (NM_001101061.1:c.2327G>A, p.Arg776His), with a heterozygous genotype in the affected calf that was not detected in any of the control genomes

including both parents. This C-to-T transition at position 20146737 on chromosome 3 located in exon 12 of *ADAMTSL4* is predicted to alter the encoded amino acid of ADAMTSL4 residue 776 (NP_001094531.1:p.Arg776His) to histidine, a less polar and more hydrophobic amino acid than arginine. The affected residue represents a probably functionally important and conserved residue located in the third of seven thrombospondin type 1 repeat domains (Figure 3). Although the arginine to histidine substitution was predicted to be neutral using PROVEAN software (Choi & Chan, 2015), it remains unclear whether this amino acid substitution affects protein folding or function. Sanger sequencing confirmed the mutant allele to be present in a heterozygous state in the affected calf and absent from its parents, given that the variant was not found in the paternal germline DNA analyzed (Figure 3). This means that the c.2327G>A variant probably arose *de novo* spontaneously during very early development of the calf.

The formal possibility exists, however, that the detected protein-changing *de novo* mutation is simply a functionally neutral change, but we regard this possibility as unlikely because of the following reasons. Various ADAMTS (a disintegrin-like and metalloproteinase with thrombospondin type 1 motifs) proteins are necessary for normal ocular development and eye function (Mead & Apte, 2018). ADAMTS-like proteins lack a metalloprotease domain, reside in the extracellular matrix and have regulatory roles and ADAMTSL4 has been implicated in fibrillin microfibril biogenesis (Gabriel et al., 2012). *ADAMTSL4* is widely expressed in non-ocular tissues as well as in various eye components, particularly in the lens equatorial epithelium

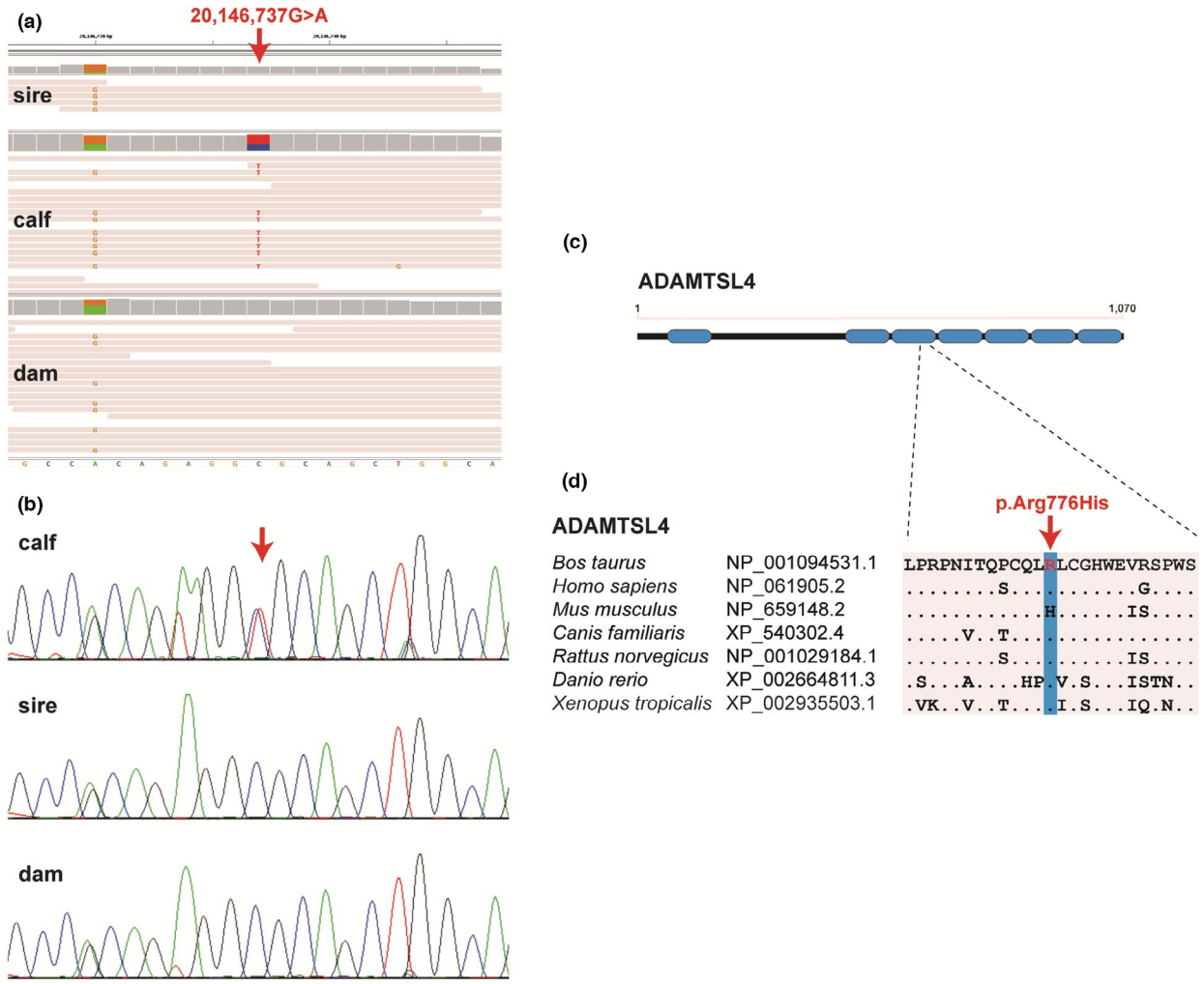


FIGURE 3 Details of the detected *ADAMTSL4* variant. (a) Integrative Genomics Viewer screenshot presenting the heterozygous single nucleotide variant (red arrow) present only in the affected calf. (b) Sanger sequencing results confirmed that the variant occurred *de novo* as sequencing of PCR products from DNA of both parents (for the sire semen) showed that the c.2327G>A variant was absent. (c) Schematic diagram of the bovine ADAMTSL4 protein that has a repetitive domain structure with seven thrombospondin type 1 repeat domains (blue). The p.Arg776His variant (red arrow) affects the third repeat. (d) Multiple species amino acid alignments encompassing the region of the variant demonstrates a high evolutionary conservation across species. The observed variant is indicated by an arrow and the respective position highlighted in gray

when the zonule attaches (Chandra et al., 2013; Gabriel et al., 2012). Besides focal retinal pigment epithelium defects, homozygous disruption of murine *Adamtsl4* resulted in a defect in the anchoring of zonule fibers to the lens surface, causing ectopia lentis, confirming its role in zonule formation (Collin et al., 2015). In humans, recessive isolated ectopia lentis (subluxation or dislocation of the human crystalline lens) and ectopia lentis et pupillae are caused by *ADAMTSL4* loss-of-functions variants (OMIM 610113). Heterozygous carriers of the known nonsense or splice site variants were reported to be apparently normal probably due to nonsense-mediated decay of the aberrant transcripts. Finally, visual inspection for large structural variants in the genome, performed after plotting the average read depth across the entire genome for the sequenced trio, revealed no obvious evidence of this type of mutation either at the chromosomal level or in the region of the *ADAMTSL4* gene (Figure S1).

Therefore, we propose the c.2327G>A variant as candidate causative variant for the observed congenital cataract phenotype based on the following arguments: (1) given that only one protein-changing *de novo* mutation event per generation is expected on average (Heinzen et al., 2015), this strongly supports the causality of the variant; (2) evolutionary conservation and expansion of ADAMTSL proteins in mammals indicates a crucial role in embryonic development (Mead & Apte, 2018); (3) in mice, it was shown that *ADAMTSL4* is strongly expressed in the lens epithelium at the lens equator throughout embryonic development (Collin et al., 2015); and (4) patients with *ADAMTSL4*-related ectopia lentis commonly present with a marked loss in visual acuity in addition to a number of possibly accompanying ocular complications including early cataract development (Ahram et al., 2009; Christensen et al., 2010). Given that this is a single case investigation and that we have no functional confirmation, this result must be considered preliminary and should be interpreted with caution. However, it must also be emphasized that the analysis was not suitable for identifying larger structural variants. Further isolated cases of cataract in cattle could be investigated for *ADAMTSL4* variants by DNA sequencing.

ACKNOWLEDGEMENTS

The authors would like to thank Cécile Meili (Braunvieh Schweiz) for contributing with information and contact to the farmer. The authors also wish to thank Swissgenetics for providing semen and Hubert Pausch, who made the sequencing data from the sire publicly available. Furthermore, we are thankful to Nathalie Besuchet-Schmutz and Simon Pot for expert assistance. We also acknowledge all bovine researchers contributing to the 1000 Bull Genomes Project and those who deposited cattle whole-genome sequencing data into public databases. This study was partially supported by the Swiss National Science Foundation, grant number 172911.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

The whole genome sequencing data of the animals generated in this study is publicly available at ENA project accession PRJEB18113 (https://www.ebi.ac.uk/ena/browser/view/PRJEB18113) with sample accessions SAMEA5159837 (case) and SAMEA6528886 (dam). The whole genome sequencing data of the sire was publicly available (ENA project accession PRJEB28191 (https://www.ebi.ac.uk/ena/browser/view/PRJEB28191), sample accession SAMEA5059753).

ORCID

Irene M. Häfliger https://orcid.org/0000-0002-5648-963X
Udo Hetzel https://orcid.org/0000-0001-9142-560X
Franz R. Seefried https://orcid.org/0000-0003-4396-2747
Cord Drögemüller https://orcid.org/0000-0001-9773-522X

REFERENCES

Ahram, D., Sato, T.S., Kohilan, A., Tayeh, M., Chen, S., Leal, S. et al. (2009) A homozygous mutation in ADAMTSL4 causes autosomal-recessive isolated ectopia lentis. *American Journal of Human Genetics*, 84, 274–278.

Bistner, S.I., Rubin, L.F. & Saunders, L.Z. (1970) The ocular lesions of bovine viral diarrhea-mucosal disease. *Pathologia Veterinaria*, 7, 275–286.

Braun, M., Struck, A.K., Reinartz, S., Heppelmann, M., Rehage, J., Eule, J.C. et al. (2019) Study of congenital Morgagnian cataracts in Holstein calves. *PLoS One*, 14, e0226823.

Chandra, A., Jones, M., Cottrill, P., Eastlake, K., Limb, G.A. & Charteris, D.G. (2013) Gene expression and protein distribution of ADAMTSL-4 in human iris, choroid and retina. *British Journal of Ophthalmology*, 97, 1208–1212.

Choi, Y. & Chan, A.P. (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, 31, 2745–2747.

Christensen, A.E., Fiskerstrand, T., Knappskog, P.M., Boman, H. & Rødahl, E. (2010) A novel ADAMTSL4 mutation in autosomal recessive ectopia lentis et pupillae. *Investigative Ophthalmology & Visual Science*, 51, 6369–6373.

Collin, G.B., Hubmacher, D., Charette, J.R., Hicks, W.L., Stone, L., Yu, M. et al. (2015) Disruption of murine Adamtsl4 results in zonular fiber detachment from the lens and in retinal pigment epithelium dedifferentiation. *Human Molecular Genetics*, 24, 6958–6974.

Detlefson, J.A. & Yapp, W.W. (1920) The inheritance of congenital cataract in cattle. *The American Naturalist*, 54, 277–280.

Gabriel, L.A., Wang, L.W., Bader, H., Ho, J.C., Majors, A.K., Hollyfield, J.G. et al. (2012) ADAMTSL4, a secreted glycoprotein widely distributed in the eye, binds fibrillin-1 microfibrils and accelerates microfibril biogenesis. *Investigative Ophthalmology & Visual Science*, 53, 461–469.

Gregory, P.W., Mead, S.W. & Regan, W.M. (1943) A congenital hereditary eye defect of cattle. *Journal of Heredity*, 34, 125–128.

Häfliger, I.M., Wiedemar, N., Švara, T., Starič, J., Cociancich, V., Šest, K. et al. (2020) Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Animal Genetics*, 51, 382–390.

Hayes, B.J. & Daetwyler, H.D. (2019) 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual Review of Animal Biosciences*, 7, 89–102.

Heinzen, E.L., Neale, B.M., Traynelis, S.F., Allen, A.S. & Goldstein, D.B. (2015) The genetics of neuropsychiatric diseases: looking in and beyond the exome. *Annual Review of Neuroscience*, 38, 47–68.

Hollmann, A.K., Dammann, I., Wemheuer, W.M., Wemheuer, W.E., Chilla, A., Tipold, A. et al. (2017) Morgagnian cataract resulting from a naturally occurring nonsense mutation elucidates a role of CPAMD8 in mammalian lens development. *PLoS One*, 12, e0180665.

Krump, L., O'Grady, L., Lorenz, I. & Grimes, T. (2014) Congenital cataracts in an Ayrshire herd: a herd case report. *Irish Veterinary Journal*, 67, 2.

Levin, L.A., Nilsson, S.F.E., Ver Hoeve, J., Wu, S., Kaufman, P.L. & Alm, A. (2011) *Adler's physiology of the eye e-book*, 11th edition. Amsterdam, The Netherlands: Elsevier Health Sciences. ISBN 978-0-323-08116-0.0.

Mead, T.J. & Apte, S.S. (2018) ADAMTS proteins in human disorders. *Matrix Biology : Journal of the International Society for Matrix Biology*, 71-72, 225–239. https://doi.org/10.1016/j.matbio.2018.06.002

Murgiano, L., Jagannathan, V., Calderoni, V., Joechler, M., Gentile, A. & Drögemüller, C. (2014) Looking the cow in the eye: deletion in the NID1 gene is associated with recessive inherited cataract in Romagnola cattle. *PLoS One*, 9, e110628.

Osinchuk, S., Petrie, L., Leis, M., Schumann, F., Bauer, B., Sandmeyer, L. et al. (2017) Congenital nuclear cataracts in a Holstein dairy herd. *Canadian Veterinary Journal*, 58, 488–492.

Schnorr, B. & Kressin, M. (2006) *Embryologie der Haustiere: Ein Kurzlehrbuch*. Leipzig: Georg Thieme Verlag.

Shiels, A. & Hejtmancik, J.F. (2017) Mutations and mechanisms in congenital and age-related cataracts. *Experimental Eye Research*, 156, 95–102.

Siepker, C.L., Zimmer, J.L., Bedard, K.M., Hart, K.A., Czerwinski, S.L. & Carmichael, K.P. (2021) Congenital cataracts and microphakia with retinal dysplasia and optic nerve hypoplasia in a calf. *Case Reports in Veterinary Medicine*, 2021, 2064103.

Williams, D.L. (2010) Congenital abnormalities in production animals. *The Veterinary Clinics of North America. Food Animal Practice*, 26, 477–486.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Fig S1

How to cite this article: Häfliger, I.M., Wolf-Hofstetter, S., Casola, C., Hetzel, U., Seefried, F.R. & Drögemüller, C. (2022) A de novo variant in the bovine *ADAMTSL4* gene in an Original Braunvieh calf with congenital cataract. *Animal Genetics*, 00, 1–6. Available from: https://doi.org/10.1111/age.13178

CNGB3 missense variant causes recessive achromatopsia in Original Braunvieh cattle

Journal:	International Journal of Molecular Sciences
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, SNP data investigation, genetic figures, co-writing original draft, writing – review and editing
Displayed version:	published version
DOI:	10.3390/ijms222212440

International Journal of
Molecular Sciences

MDPI

Article

CNGB3 Missense Variant Causes Recessive Achromatopsia in Original Braunvieh Cattle

Irene M. Häfliger [1,†], Emma Marchionatti [2,†], Michele Stengård [3,†], Sonja Wolf-Hofstetter [1], Julia M. Paris [1], Joana G. P. Jacinto [4], Christine Watté [3], Katrin Voelter [5], Laurence M. Occelli [6], András M. Komáromy [6], Anna Oevermann [7], Christine Goepfert [8], Angelica Borgo [9], Raphaël Roduit [9], Mirjam Spengeler [10], Franz R. Seefried [10] and Cord Drögemüller [1,*]

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch (I.M.H.); hofstettersonja@hotmail.com (S.W.-H.); julia-1991@hotmail.com (J.M.P.)
2 Clinic for Ruminants, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland; emma.marchionatti@vetsuisse.unibe.ch
3 Division of Ophthalmology, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Länggassstrasse 128, 3001 Bern, Switzerland; michele.stengard@vetsuisse.unibe.ch (M.S.); christine.watte@vetsuisse.unibe.ch (C.W.)
4 Department of Veterinary Medical Sciences, University of Bologna, 50 Ozzano Emilia, 40064 Bologna, Italy; joana.goncalves2@studio.unibo.it
5 Ophthalmology Section, Vetsuisse Faculty, University of Zurich, Winterthurerstrasse 260, 8057 Zurich, Switzerland; kvoelter@vetclinics.uzh.ch
6 College of Veterinary Medicine, Michigan State University, 736 Wilson Rd., East Lansing, MI 48824, USA; occelli@msu.edu (L.M.O.); komaromy@msu.edu (A.M.K.)
7 Division of Neurological Sciences, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland; anna.oevermann@vetsuisse.unibe.ch
8 COMPATH, Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Länggassstrasse 122, 3001 Bern, Switzerland; christine.goepfert@vetsuisse.unibe.ch
9 Department of Ophthalmology, University of Lausanne, Jules-Gonin Eye Hospital, Av. de France 15, 1004 Lausanne, Switzerland; angelica.borgo@fa2.ch (A.B.); raphael.roduit@fa2.ch (R.R.)
10 Qualitas AG, Chamerstrasse 56, 6300 Zug, Switzerland; mirjam.spengeler@qualitasag.ch (M.S.); franz.seefried@qualitasag.ch (F.R.S.)
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-(0)31-684-25-29
† Contributed equally.

Abstract: Sporadic occurrence of inherited eye disorders has been reported in cattle but so far pathogenic variants were found only for rare forms of cataract but not for retinopathies. The aim of this study was to characterize the phenotype and the genetic aetiology of a recessive form of congenital day-blindness observed in several cases of purebred Original Braunvieh cattle. Electroretinography in an affected calf revealed absent cone-mediated function, whereas the rods continue to function normally. Brain areas involved in vision were morphologically normal. When targeting cones by immunofluorescence, a decrease in cone number and an accumulation of beta subunits of cone cyclic-nucleotide gated channel (CNGB3) in the outer plexiform layer of affected animals was obvious. Achromatopsia is a monogenic Mendelian disease characterized by the loss of cone photoreceptor function resulting in day-blindness, total color-blindness, and decreased central visual acuity. After SNP genotyping and subsequent homozygosity mapping with twelve affected cattle, we performed whole-genome sequencing and variant calling of three cases. We identified a single missense variant in the bovine *CNGB3* gene situated in a ~2.5 Mb homozygous genome region on chromosome 14 shared between all cases. All affected cattle were homozygous carriers of the p.Asp251Asn mutation that was predicted to be deleterious, affecting an evolutionary conserved residue. In conclusion, we have evidence for the occurrence of a breed-specific novel *CNGB3*-related form of recessively inherited achromatopsia in Original Braunvieh cattle which we have designated OH1 showing an allele frequency of the deleterious allele of ~8%. The identification of carriers will enable selection against this inherited disorder. The studied cattle might serve as an animal model to further elucidate the function of *CNGB3* in mammals.

Citation: Häfliger, I.M.; Marchionatti, E.; Stengård, M.; Wolf-Hofstetter, S.; Paris, J.M.; Jacinto, J.G.P.; Watté, C.; Voelter, K.; Occelli, L.M.; Komáromy, A.M.; et al. *CNGB3* Missense Variant Causes Recessive Achromatopsia in Original Braunvieh Cattle. *Int. J. Mol. Sci.* **2021**, *22*, 12440. https://doi.org/10.3390/ijms222212440

Academic Editor: Akiko Maeda

Received: 29 October 2021
Accepted: 17 November 2021
Published: 18 November 2021

Keywords: *Bos taurus*; animal model; day-blindness; retina; development; mendelian genetics; rare disease; precision medicine

1. Introduction

Food animal ophthalmology is a neglected area of veterinary medicine [1,2]. In farm animal practice, conditions such as visual impairment or even blindness, despite having a negative impact on behavior and welfare, are rarely considered, as there is usually no profoundly damaging economic impact on animal production [3].

Sporadic occurrence of inherited eye disorders in livestock species such as cattle has been reported [4–6]. However, many affected newborns go unreported or escape surveillance systems [4]. In addition to hereditary reasons, environmental causes such as vitamin A deficiency and bacterial or viral infections must also be taken into account [5–7]. Although congenital eye defects are rare, they are important and should be considered because they often follow monogenic recessive inheritance [4,6]. Therefore, such inherited disorders can rapidly gain in prevalence due to the undetected use of carriers, especially in the course of artificial insemination in cattle. So far pathogenic variants were found only for rare breed-specific recessive forms of cataract in Romagnola (OMIA 001936-9913) and Holstein (OMIA 002111-9913) cattle [8,9]. Such findings enable selection against these disorders within the affected populations. Furthermore, the *NID1*-related cataract observed in Romagnola was the first report of a naturally occurring mutation leading to a non-syndromic form of cataract in a mammalian species [8]. Therefore, the genetic study performed in cattle added the affected gene to the list of candidate genes for inherited forms of nuclear cataract in humans, illustrating the impact of studying eye conditions in domestic animals.

Achromatopsia, an inherited retinal disease characterized by the loss of cone photoreceptor function resulting in day-blindness, total color-blindness, and decreased central visual acuity, has not yet been described in cattle. Pathogenic variants in six genes (*CNGA3*, *CNGB3*, *GNAT2*, *ATF6*, *PDE6C*, and *PDE6H*) have been identified in humans with achromatopsia (OMIM 216900) [10]. In sheep, a form of *CNGA3*-related achromatopsia has been characterized (OMIA 001481-9940) [11]. This ovine condition was intensively used for functional restoration of cone function [12], highlighting the biomedical value of such large animal models [13,14] in addition to the direct benefits for animal breeding and animal health.

The aim of this study was to characterize the phenotype and the underlying causative genetic defect for this presumably new form of congenital day-blindness observed in several cases of purebred Original Braunvieh cattle from Switzerland.

2. Results

2.1. Clinical Description

Initially, in 2017, a four-month-old calf (case 2) was presented to the University of Zurich for evaluation of suspected day blindness. Ocular examination revealed bilaterally absent menace responses and dazzle reflexes. Pupillary responses were positive. Chromatic light stimulation (Melan) revealed bilaterally minimal dazzle reflexes with red and blue light, pupillary responses were normal with blue and reduced with red light stimulation. The remainder of the ocular exam (slit-lamp biomicroscopy, indirect and direct ophthalmoscopy, intraocular pressure) was within normal limits with the optic nerve head showing a deep physiologic cup. The calf was able to complete an obstacle course in dim light, but not in bright light.

In 2020, a 5.5-month-old calf (case 12) was presented to the University of Bern for evaluation of a suspected vision disturbance. The farmer noted that the calf was hesitant to walk and often collided with obstacles in its environment, especially when separated from the mother (Video S1). Observation of the calf revealed extremely poor navigation

of large and small objects in the examination hall in ambient room light. Navigation of obstacles and recognition of large high-contrast obstacles after room lights were turned off was slightly improved. Ophthalmic examination revealed absent menace response in both eyes (OU) in dim, ambient and bright (outdoor) light conditions. The direct and consensual pupillary light responses as well as dazzle reflexes to white light were reduced in both eyes. A slightly reduced pupillary light response was noted with bright blue light stimulation and severely reduced to bright red light stimulation. Slit-lamp biomicroscopic examination of the eyelids, conjunctiva, cornea, anterior chamber, lens and anterior vitreous was normal in both eyes with the exception of mild mydriasis OU. The pupils were symmetrical OU. Fluorescein staining of the cornea was negative OU. Indirect ophthalmoscopy revealed normal appearance to the optic nerve, tapetal and non-tapetal retina, and retinal vessels bilaterally. Intraocular pressure was measured within normal limits (11 mmHg OD, 14 mmHg OS).

Electroretinography in the affected calf revealed severely reduced light-adapted single cone and cone flicker responses, while the dark-adapted mixed cone-rod response was considered normal (Figure 1). Combined with the behavioral observations and clinical findings, the ERGs support an achromatopsia diagnosis based on the specific loss of cone-mediated retinal function without any retinal degenerative changes.

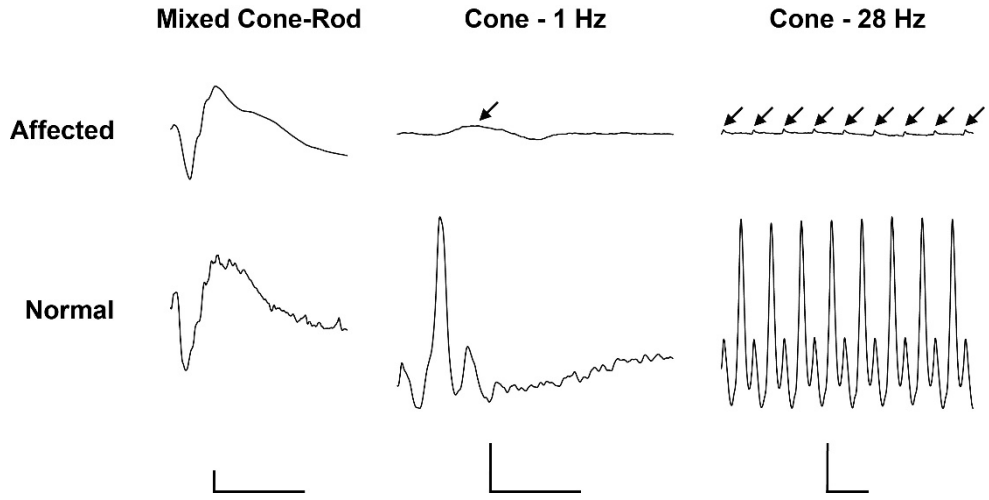


Figure 1. Electroretinogramm (ERG) of a 5.5-month old achromatopsia-affected Original Braunvieh (case 12) and a seven-month old Hereford control cattle. Dark-adaptation responses in representative normal and affected calves. While the dark-adapted mixed cone-rod responses were comparable between the two animals, the light-adapted single cone (1 Hz) and cone flicker (28 Hz) responses were severely reduced (arrows). Calibration bars: vertical = 100 μV, horizontal = 50 ms.

2.2. Pathological Phenotype

During necropsy of case 12 no gross lesions were detectable and brain areas involved in vision were morphologically normal.

The retina were targeted by immunofluorescence. We observed the expression of CNGB3 in the outer segment (OS) of cones in the control animal, while this expression was decreased in the affected animal. We also detected a significant difference of staining in the outer nuclear layer (ONL), where the CNGB3 protein seemed to be concentrated in the affected animal. This may suggest a mislocalization of the mutated CNGB3. Immunoreactivity of CNGB3 in the outer plexiform layer (OPL) may be due to an unspecific signal as the localization is observed in both control and affected animals. (Figure 2a). Further experiments on other affected calves (we analysed only one) or by transfection of the mutated CNGB3 in cells will confirm the potential higher stability of the mutated cyclic nucleotide-gated channel.

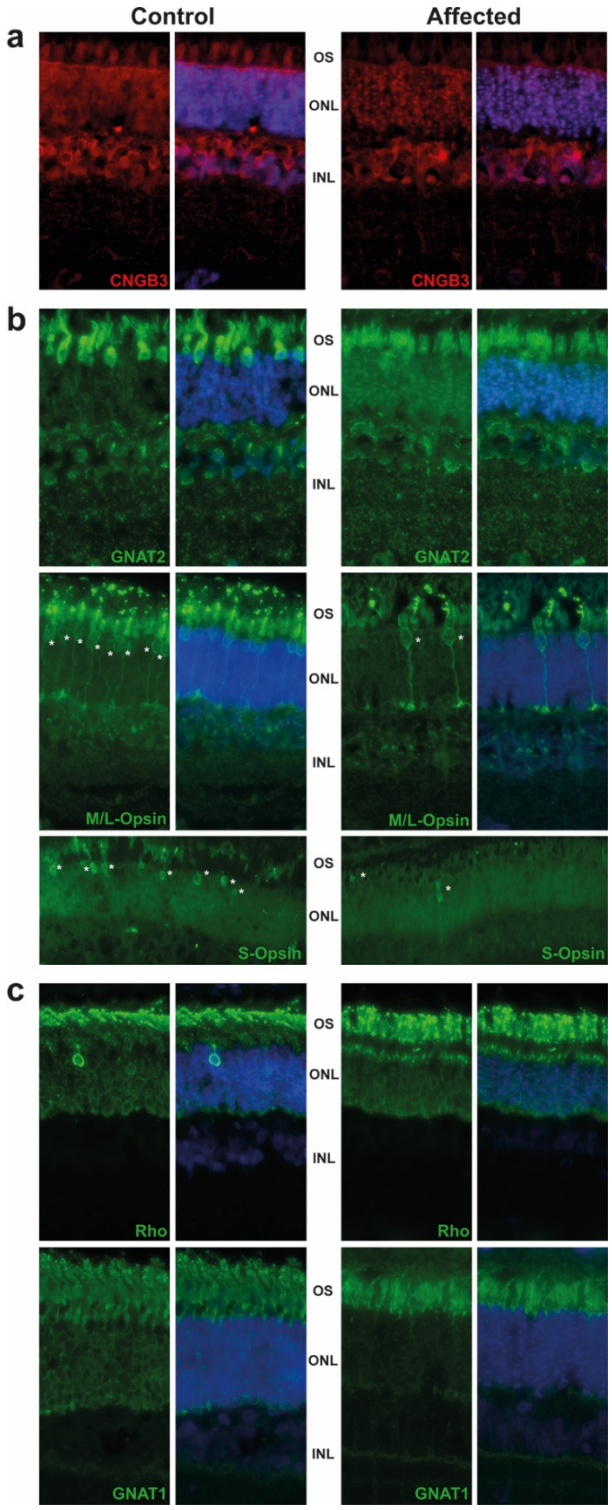


Figure 2. Immunostaining of retinal markers in a five-month old achromatopsia-affected cattle (case 12) and a control cattle of the same age. Cyclic nucleotide gated channel subunit beta 3 (CNGB3) (**a**), Cones markers (GNAT2, M/L-OPSIN, S-OPSIN) (**b**) and rods markers (GNAT1, RHODOPSIN) (**c**) are immunostained in both control and affected animals accordingly to conditions described in Table S2. Cell nuclei are shown in blue with DAPI. Images acquired at equal distances from the optic nerve head for each protein. Asterisks (*) show M/L- and S-Opsin cones present in both control and affected animal. Negative controls without primary antibody were performed (not shown).

Based on GNAT2 and cone opsins staining, we clearly observed a decrease in the cone OS, which are shorter and abnormally shaped. Although we did not count the number of cones per retina, we observed a decrease in their number, in affected animal in comparison to control, as shown in M/L-and S-Opsin staining (Figure 2b, white asterisk). These results are in correlation with the ERG and are consistent with the progressive loss of cone outer segments seen in other species with achromatopsia. We also observed a slight decrease of the rod OS length in affected animals and a mislocalization of Rho (Figure 2c).

2.3. Pedigree Analysis

The initially studied cases 1 and 2 were both the only affected animals in two different Swiss herds of purebred Original Braunvieh cattle. The sire of these two cases was a natural service purebred Original Braunvieh bull, which sired a further 24 apparently normal offspring within two years. A query to the Original Braunvieh breeders in Switzerland revealed further evidence of ten similar cases sired by different bulls collected over a period of three years. The available pedigree records of all 12 cases were analyzed and multiple inbreeding loops between the parents were found (Figure 3). We detected a single common ancestor occurring 8–11 generations ago. Due to the obvious history of inbreeding, a recessive inherited condition was considered. In light of the obvious consanguinity as well as the apparently unaffected parents, we hypothesize that the achromatopsia-affected calves might be explained by a recessively inherited variant. The founding mutation thus probably occurred many generations before the cases occurred. The causal variant was probably spread by the common ancestor, an artificial insemination bull born in 1961, as well as by some of his male descendants.

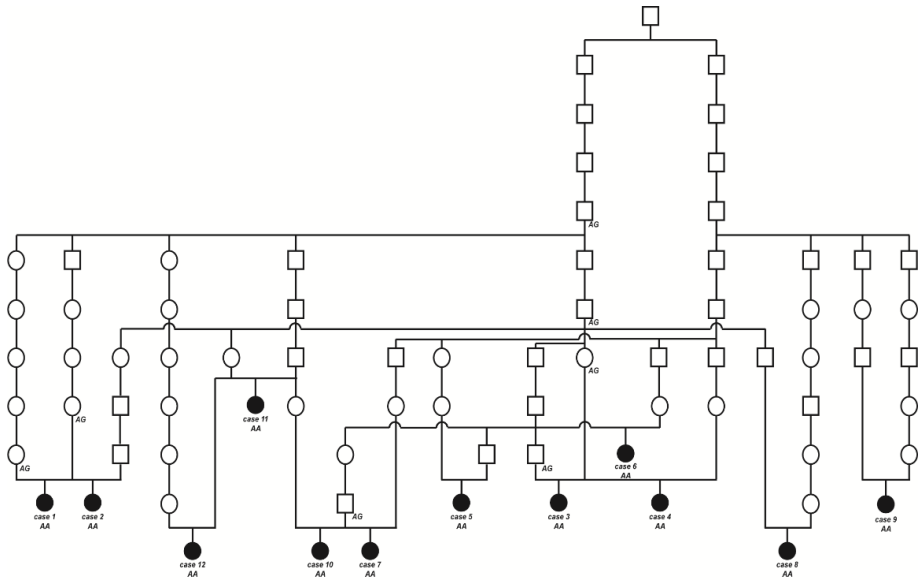


Figure 3. Pedigree of the Original Braunvieh cattle family segregating for achromatopsia suggested monogenic autosomal recessive inheritance. Affected animals are filled symbols. Open symbols represent normal cattle. DNA samples were available from animals with genotypes for the *CNGB3* XM_015474554.2:c.751G>A variant are given below the symbols.

2.4. Genetic Analysis

SNP genotyping data for twelve affected cattle identified two shared ROHs between all cases on chromosome 11 and 14 (Figure 4a). On chromosome 11 all animals were homozygous for 101 SNP markers from 66,668,989 to 68,938,216 corresponding to a repeatedly detected strong selection signature of the Original Braunvieh breed encompassing a genome region with 24 protein-coding genes [15–17]. As no candidate gene for a retinopathy was contained in that region, we focused on the second ROH found on chromosome 14. All twelve affected cattle were homozygous at 28 SNP markers on chromosome 14 from

74,306,245 to 76,800,429, which allowed the identification of a single disease-associated IBD haplotype shared by all cases, limiting the critical region to 2,494,184 bp on chromosome 14 (Table S1; Figure 4a,b). Interestingly, the bovine homolog of *CNGB3*, a gene that causes achromatopsia in other species, maps to that genome region at 76 Mb (Figure 4c).

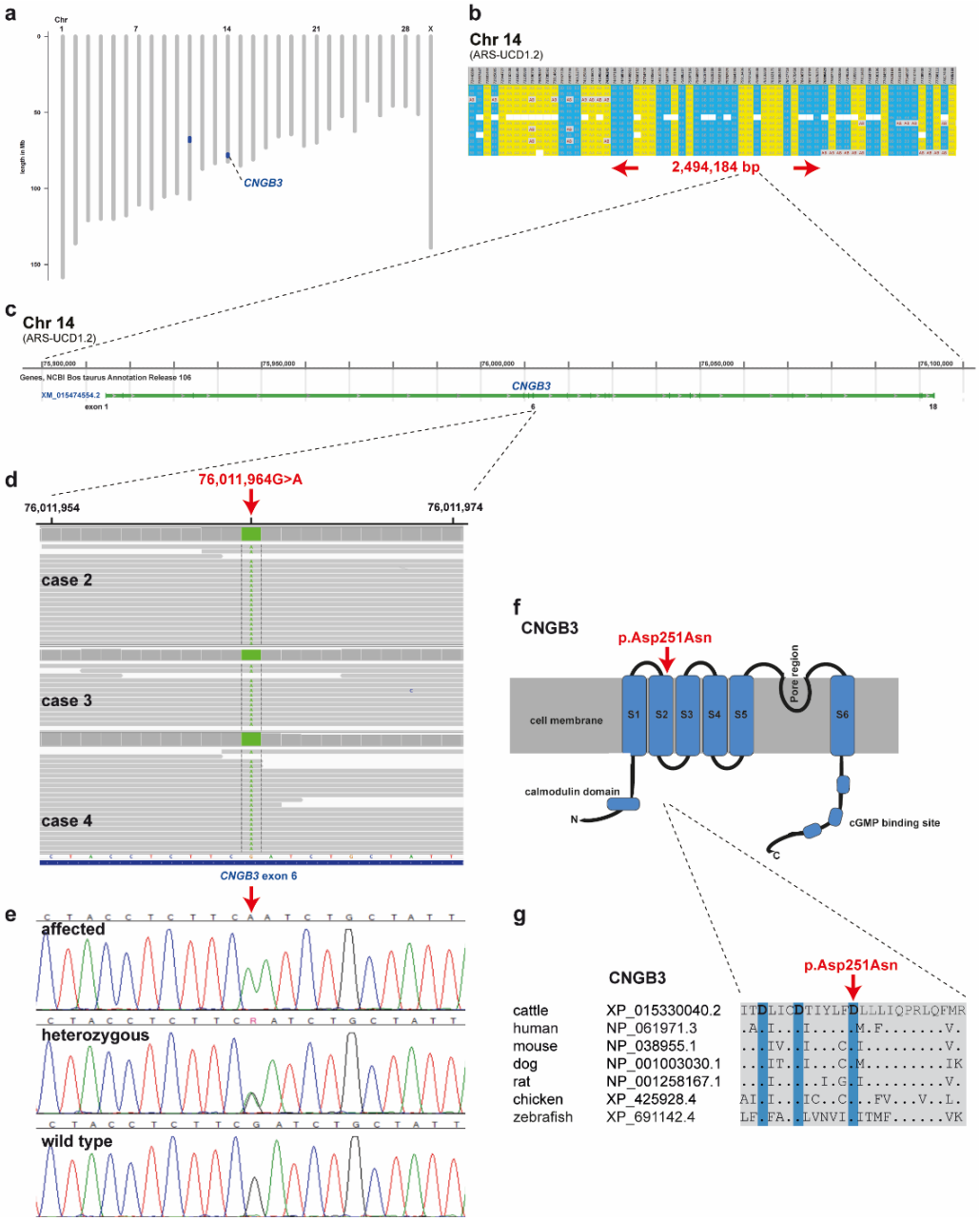


Figure 4. Achromatopsia-associated *CNGB3* missense variant in Original Braunvieh. (**a**) Genome-wide homozygosity mapping presenting the two homozygous blocks shared in 12 affected calves in blue. Note the red arrow highlighting the *CNGB3* gene on cattle chromosome 14. (**b**) Schematic representing the SNP genotypes of 12 affected calves on chromosome 14. Each horizontal lane represents one calf with yellow and blue shading, indicating shared homozygosity. Grey shading indicates a heterozygous genotype and white indicates missing genotypes. The genome positions of markers are indicated above the figure. The red arrows indicate the consensus homozygous region that spans approximately 2.5 Mb. (**c**) *CNGB3* gene structure showing the location of the exon 6 variant. (**d**) Genome viewer screenshot presenting the homozygous Chr14: g.76011964A>G variant in three affected calves. (**e**) Electropherograms showing the different genotypes identified via Sanger sequencing. (**f**) Localization of the missense variant (red arrow) with respect to the topological model of the CNGB3 protein. (**g**) Across species sequence alignment of the affected CNGB3 S2 domain. Note that the missense variant (red arrow) affects the evolutionary conserved Tri-Asp motif that is highlighted in blue.

Subsequently, we sequenced the genomes of three of the affected cattle (cases 2, 3 and 4) and searched for private variants that were exclusively present in a homozygous state in all three affected cattle and absent or only heterozygous in the genomes of 567 other cattle. Beside 34 non-coding variants, all located in the critical region on chromosome 14, this analysis identified a single homozygous private protein-changing variant in *CNGB3*, a known candidate gene for achromatopsia. The variant can be designated as chr14: 76011964G>A (ARS-UCD1.2 assembly) (Figure 4d). It is a missense variant, XM_015474554.2: c.751G>A, predicted to change a highly conserved aspartic acid residue in the second S2 domain of CNGB3, XP_015330040.2: p.Asp251Asn (Figure 4f,g). In silico analysis predicted the functional effect of p.Asp251Asn as deleterious using PROVEAN software (score −4.985) [18].

We confirmed the presence of the *CNGB3* missense variant by Sanger sequencing (Figure 4e). The genotypes at the variant co-segregated with the achromatopsia phenotype as expected for a monogenic autosomal recessive mode of inheritance (Figure 3). All twelve available DNA samples from the achromatopsia-affected cattle carried the mutant allele in a homozygous state, while their parents were heterozygous, as expected for obligate carriers (Figure 3; Table 1).

Table 1. Association of the missense variant in *CNGB3* with the achromatopsia phenotype in Original Braunvieh cattle.

	GG	AG	AA
Achromatopsia-affected calves			12
Obligate carriers [a]		5	
Other Original Braunvieh cattle [b,c]	2477	463	12
Brown Swiss cattle [c]	14,976	52	
Holstein cattle [c]	14,825		
Simmental cattle [c]	2021		
Sequenced cattle genomes from various breeds (local Swiss cohort) [d]	552	15	
Control cattle from various breeds (1000 Bull Genomes project) [e]	3298	7 [f]	1 [g]

[a] parents of affected animals were classified as obligate carriers. [b] phenotypes are unknown. [c] Axiom® genotype data from population-wide routine genomic testing. [d] 567 genomes of the Swiss Comparative Bovine Resequencing project including 92 Original Braunvieh cattle. [e] run 8: 3306 genomes including 58 Original Braunvieh cattle. [f] exclusively Original Braunvieh. [g] case 2 was added to the 1000 Bull Genomes project.

We also genotyped the *CNGB3*: c.751G>A variant in a population control cohort comprising 2952 Original Braunvieh cattle without any phenotypic records. The mutant *CNGB3* allele was detected in the homozygous state in 12 of the cattle, whereas 463 were heterozygous carriers revealing an allele frequency of the mutant allele of 8.2% (Table 1). Interestingly, the mutant allele was absent from more than 35,000 cattle of various other breeds (Table 1). We found some rare heterozygous carriers with an allele frequency of 0.2% only in the Brown Swiss population of Switzerland (Table 1).

3. Discussion

To date, no genetic mutations have been associated with retinopathies in cattle. Affected Original Braunvieh calves with suspected vision disturbance suffer from day-blindness due to congenitally reduced cone-mediated function of the retina. Our clinicopathological evaluation of affected calves supported a diagnosis of achromatopsia based on abnormally appearing cone outer segments and normal rod photoreceptors. Furthermore, neither retinal degenerative changes nor abnormalities in the central visual pathways were observed. Pedigree and ROH analysis suggested an autosomal recessive mode of

inheritance. Genome-wide homozygosity mapping using SNP array data was used successfully for high-resolution mapping of two critical regions of shared homozygosity. We performed whole-genome sequencing on three affected Original Braunvieh calves with day-blindness to identify variants associated with the phenotype. The similar clinical presentation between familial achromatopsia in humans and bovine recessive day-blindness led to the hypothesis that a protein-changing variant within *CNGA3*, *CNGB3*, *GNAT2*, *ATF6*, *PDE6C*, and *PDE6H* would be associated with achromatopsia of Original Braunvieh calves. Whereas only *CNGB3* was located in an IBD segment, we identified a missense variant in *CNGB3*: c.751G>A, p.Asp251Asn that significantly associated with the phenotype. Nine additional affected calves were subsequently genotyped and homozygous for the missense variant. Therefore, the cattle studied could serve as an animal model to further investigate the function of *CNGB3* in mammals.

In domestic animals, to this point in sheep [11–13] and dogs [19–22] (previously reported as cone degeneration and canine hemeralopia), the underlying genetics of different forms of achromatopsia are reported. Cones alone are affected in Alaskan Malamute (OMIA 001365-9615) and the German shorthaired pointers (OMIA 001676-9615) because of breed specific mutations in *CNGB3*, a cone–specific gene. Cone cyclic nucleotide-gated channels (CNG) are tetramers formed by three CNGA3 and one CNGB3 subunit; CNGA3 subunits can function as homotetrameric channels but CNGB3 exhibits channel function only when co-expressed with CNGA3 [23]. A 140-kb deletion and a missense mutation in *CNGB3* occurs in achromatopsia-affected dogs of multiple breeds [20]. Interestingly, the described canine missense variant also leads to an exchange of an aspartic acid with an asparagine residue, compromising a critical functional domain, and the phenotype seen in homozygous dogs represents a loss of function. Similar to what we found in Original Braunvieh cattle, the *CNGB3* missense mutation causing achromatopsia in German shorthaired pointers is also located in exon 6 (c.784G, p.Asp262Asn), affecting the corresponding residue of a conserved region of the same gene, suggesting an important role for this aspartate residue in channel biogenesis and/or function [21]. The flanking region surrounding these missense mutations is well conserved between species and is predicted to encode the second transmembrane domain of the CNGB3 protein containing three Asp residues designated the *tri-Asp motif* and conserved in all CNG channels [21]. Mutations of these conserved aspartate residues result in the absence of nucleotide-activated currents in heterologous expression. Aspartate is a negatively charged, polar amino acid found in both dogs and cattle with achromatopsia replaced by asparagine (Asn), another polar amino acid, which differs only in that it contains an amino group in place of one of the oxygens found in aspartate (Asp) and thus lacks a negative charge. Obviously, retinopathies associated with missense mutations draw attention to amino acids important for understanding the structure-function properties of functionally important channels. By in vitro follow-up studies of *CNGB3*-related canine achromatopsia it was found that Asp/Asn mutations affect the heteromeric subunit assembly of the six transmembrane-spanning helices (S1–S6), resulting in the loss of these inter-helical interactions altering the electrostatic equilibrium within in the S1–S4 bundle [21]. Although disease-causing variants within the S2 segment of human CNGB3 have not been reported (OMIM 605080), a study involving a missense mutation p.Asp211Glu at S2 of CNGA3 confirmed that variations in a conserved region could lead to cone dysfunction [24].

In Switzerland, the Original Braunvieh population is the ancestor of the world-renowned Brown Swiss population, which originated in North America from animals obtained in Switzerland at the turn of the century around 1900 [25]. Therefore, we speculate that the sporadic occurrence of *CNGB3*-carriers in the current Brown Swiss population indicates that the mutation might have arisen before that time and predates modern pedigree records. In recent decades, outbreaks of four undesirable genetic defects (weaver disease, spinal dysmyelination, spinal muscular atrophy, and arachnomelia) have occurred in Brown Swiss cattle. This report represents the first genetic disorder known in Original

Braunvieh cattle which we have designated OH1, and the obtained results enable targeted selection to avoid the occurrence of further affected animals in future.

In summary, this study highlights the strong genetic similarities between human and bovine achromatopsia, suggesting that bovine achromatopsia, similar to that found in dogs, could serve as an excellent model for developing treatment strategies for humans.

4. Materials and Methods

4.1. Animal Selection for Genetic Analysis

This study was conducted with 248 Original Braunvieh cattle samples. The case cohort of this study consisted of twelve purebred Original Braunvieh cattle with suspected congenital vision disturbance reported to the breeding organization by different farmers between 2017 and 2020 (Table S1). In addition, either hair root or EDTA blood samples of three dams and two sires were collected for the genetic analysis, and genomic DNA was extracted using the Promega Maxwell® RSC system (Promega, Dübendorf, Switzerland). The remaining 231 male Original Braunvieh cattle were used as population controls. These bulls had reliable phenotype records on normal vision because they very carefully examined by veterinarians before being used for artificial insemination. Before admission to the insemination station, these young bulls are carefully examined and these examinations include, in particular, the consideration of the presence of possible congenital disorders, including a standard ophthalmological examination of the eyes.

Once the most likely causative variant was discovered, it was added to two Swiss Axiom® genotyping arrays (Thermo Fisher Scientific, Waltham, MA, USA) routinely used for genomic selection. Thus, after two years of population-wide genotyping in Swiss dairy populations for the purpose of genomic selection, more than 30,000 genotypes for the *CNGB3* variant were available. These were mainly determined in the four largest Swiss dairy cattle populations (Brown Swiss, Holstein, Original Braunvieh and Simmental).

4.2. Ophthalmological Examination including Electroretinography

A four-month-old calf (case 2) was presented to the University of Zürich Food Animal Clinic and the Ophthalmology Section in summer 2017 for evaluation of suspected vision disturbance.

A 5.5-month-old calf (case 12) was presented to the University of Bern Food Animal Clinic and the Division of Ophthalmology in the year 2020 for evaluation of suspected vision disturbance. Pupillary light responses were examined with bright red and blue light stimulation (Melan-100, 200–250 kcd, Iris-vet series, BioMed Vision Technologies, Ames, IA, USA). Furthermore, slit-lamp biomicroscopic examination (SL-17 Portable Slit Lamp, Kowa, Japan) of the adnexa and anterior segments, fluorescein staining (Contacare Ophthalmics and Diagnostics, Gujarat, India) of the corneas and indirect ophthalmoscopy (Omega 500 Binocular Indirect Ophthalmoscope, Heine Optotechnik GmBH, Gilching, Germany) of the ocular fundi were carried out. Intraocular pressures were measured by rebound tonometry (Tonovet Rebound Tonometer, Icare, Finland).

Electroretinograms (ERGs) were recorded under general anesthesia in case 12 following clinical and behavioural examinations. A jugular intravenous catheter was placed and the calf was sedated with xylazine 0.2 mg/kg IM. Induction was performed with ketamine 4 mg/kg intravenous, and 10 min after sedation the calf was placed in lateral recumbency. The head was positioned with cushions to facilitate access to the eye for testing. Anesthesia was maintained using ketamine continuous-rate infusion 3 mg/kg/h and xylazine continuous-rate infusion 0.05 mg/kg/hr. Flow-by oxygen was administered continuously via nasal oxygen catheter. All recordings were conducted on the right eye, following dilation of the pupil with 1% tropicamide ophthalmic solution. ERGs were recorded using the RetiPORT ERG system (Roland Consult, Brandenburg an der Havel, Germany). Two platinum subdermal needle electrodes (Grass Safelead Needle electrodes, Grass Technologies, West Warwick, RI, USA) were used: The reference electrode was placed subcutaneously approximately 10 mm from the lateral canthus, and the ground electrode

was placed over the occipital protuberance. An ERG-Jet® corneal electrode (Fabrinal SA, La Chaux-de-Fonds, Switzerland) was used as the active electrode and applied with 2.5% hypromellose ophthalmic demulcent solution. Flash stimuli and light adaptation were delivered using a handheld Mini Ganzfeld (Roland Consult).

Following 20 min of dark adaptation, mixed cone-rod responses were recorded with a flash intensity of 0.096 cd.s/m^2 (average of 3 sweeps at 0.1 Hz). Subsequently, the eye was light-adapted for 5 min to a white uniform background light of 30 cd/m^2 and single cone (average of 3 sweeps at 1.0 Hz) and cone flicker (average of 3 sweeps at 28 Hz) responses were recorded with a 3.0 cd.s/m^2 flash intensity. For all recordings, the filters were set to allow a bandpass of 1 to 300 Hz.

4.3. Targeting Cones by Immunofluorescence

After slaughtering, the enucleated calf eyes of case 12 were fixed in Bouin's solution for 24 h, trimmed and paraffin embedded. The 3 µm-embedded paraffin sections were further processed for immunofluorescence. Briefly, retina sections were first deparaffinized by successive baths (three different Xylol baths of 5 min, 3 min, 3 min respectively; and six ethanol baths: from 100% to 70%; then washed several times in water). Sections were then boiled 30 min in a Dako antigen retrieval solution (Agilent S169984-2) and left to cool down for 45 min. Retina sections were incubated for 1 h in blocking solution and incubated with primary antibodies as indicated in Table S2. Following incubation with primary antibodies, sections were washed 3 times in PBS and incubated for 1.5 h at RT with the secondary antibodies (Table S2). After three successive washing steps in PBS, sections were treated for 25 min in 0.1% Sudan black B (Sigma 380B)/70% ethanol. Then, sections were washed again twice in ethanol 70%, and three times in PBS 0.02% Tween and counterstained with 49,6-Diamidino-2-phenylindole (DAPI) to identify retinal cell layers. After three washing steps in PBS, sections were mounted with antifadent citifluor solution (Electron microscopy sciences, Hatfield, PA, USA). Immunostaining was visualized under a fluorescence microscope (Leica, Switzerland). Incubation with the secondary antibody alone was used as a negative control, and every image acquisition of the retina was made at the same distance from the optic nerve head for each antigen.

4.4. Morphology and Histopathology of the Visual Pathway

The head of case 1 was taken for gross and histopathological evaluation after slaughtering at the age of five months. The visual pathways of the affected cattle were evaluated and biopsies fixed in 4% formaldehyde for routine histopathological evaluation with haematoxylin and eosin (H&E) staining of the optic nerve, optic tract, optic chiasm, lateral geniculate nucleus and the visual cortex.

4.5. SNP Genotyping and Subsequent Homozygosity Mapping with 12 Affected Cattle

Genotype data for the twelve achromatopsia-affected cattle (cases 1–12) were obtained with an Illumina BovineHD BeadChip array. The PLINK v1.9 software [26] was used to perform basic quality filtering of the dataset. For homozygosity mapping, the genotype data for the twelve affected cattle were used. Markers on the sex chromosomes were excluded. The following PLINK option parameters were applied (–homozyg-snp 10; –homozyg group; –homozyg-density 30; –homozyg-gab 1000; –homozyg-window-het 0; –homozyg-window-missing 0) to search for extended regions of homozygosity (ROH) indicating chromosomal region of identity-by-descent (IBD). ROH analyses were performed using an imputed dataset that included the entire Swiss genotype archive for Original Braunvieh. Animals were genotyped using several routinely available array chips that included between nine and 777 k SNPs. The available genotype archive was used in a two-step imputation approach and was imputed first to a density of 150 k. Subsequently, imputation to 777 k-density was carried out using 150 k data. A number of 2507 and 351 Original Braunvieh reference animals were available for 150 k and 777 k imputation, respectively. FImpute v2.2 software was used with default parameters for both steps [27]. In each step, SNPs

with a minor allele frequency (MAF) lower than 1% were removed from the dataset. The final marker set included 114891 and 681179 SNPs for each density (150 k and 777 k), respectively. SNPs were filtered using the following thresholds: MAF higher than 0.01 and an SNP call rate higher than 0.99 in the genotype data from the reference population. The output interval was displayed in Excel spreadsheets to find overlapping regions (Table S1). All positions correspond to the ARS-UCD1.2 reference genome assembly.

4.6. Whole-Genome Resequencing and Variant Calling

Three Illumina TruSeq PCR-free libraries with ~500 bp insert size were prepared from three affected cattle (cases 2, 3 and 4). We collected 2 × 150 bp paired-end reads on a NovaSeq 6000 instrument. Mapping to ARS-UCD1.2 reference genome assembly was performed as described [28]. The sequence data were deposited under study accession PRJEB28191 and sample accessions SAMEA4644768, SAMEA6528889 and SAMEA6528891 at the European Nucleotide Archive.

Variant calling including single-nucleotide variants (SNVs) and small indels was performed as described [29]. To predict the functional effects of the called variants, SnpEff software v4.3 [30] together with the ARS-UCD1.2 reference genome assembly and NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/; accessed on 30 June 2021) was used. For private variant filtering we used control genome sequences from 567 cattle of diverse breeds including 119 Original Braunvieh animals. These genomes were produced during the Swiss Comparative Bovine Resequencing project and made publicly available (https://www.ebi.ac.uk/ena/browser/view/PRJEB18113/; accessed on 30 June 2021). The most likely pathogenic missense variant in *CNGB3* was inspected for its presence in a global control cohort of 3306 genomes with a sequence depth of at least 8x from a variety of breeds including 92 Original Braunvieh animals (1000 Bull Genomes Project run 8; www.1000bullgenomes.com accessed on 30 June 2021) [29].

4.7. Genotyping Assays

Two genotyping tests were developed for the XM_015474554.2:c.751G > A missense variant in the *CNGB3* gene to confirm segregation with disease and to estimate the allele frequency in the population.

4.7.1. PCR and Sanger Sequencing

We designed a specific PCR for the targeted genotyping of the chr14:76011964G>A variant. PCR was performed for 30 cycles using Amplitaq Gold Master Mix (Thermofisher, Rotkreuz, Switzerland) in a 10 µL reaction containing 10 ng genomic DNA and 5 pmol of each primer (F 5′-CCTGTGGCTCTCACTTGTCA-3′ and R 5′- CTCCCGAGCCCCTACTTACT-3′). After treatment with exonuclease I and alkaline phosphatase, PCR amplicons were sequenced on an ABI 3730 DNA Analyzer (Thermofisher, Rotkreuz, Switzerland). Sanger sequences were analyzed using the Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA).

4.7.2. Axiom® Genotyping Array

Two fully customized Axiom® genotyping arrays (Thermo Fisher Scientific, Rotkreuz, Switzerland) designed for genomic selection purpose in Swiss dairy cattle populations designated as SWISScow (96-array layout with 314,744 markers) and SWISSLD1 (384-array layout with 64,212 markers) both included the chr14:76011964G>A variant.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms222212440/s1.

Author Contributions: Conceptualization, F.R.S. and C.D.; validation, I.M.H., E.M., M.S. (Michele Stengård) and J.M.P.; formal analysis, I.M.H., S.W.-H., M.S. (Michele Stengård), C.W., E.M., J.G.P.J., K.V., L.M.O., A.M.K., A.O., C.G., A.B., R.R. and M.S. (Mirjam Spengeler); investigation, I.M.H., S.W.-

H., E.M., M.S. (Michele Stengård) and R.R.; data curation, I.M.H.; writing—original draft preparation, I.M.H., E.M., M.S. (Michele Stengård), A.M.K., R.R. and C.D.; writing—review and editing, I.M.H., S.W.-H., J.M.P., M.S. (Michele Stengård), C.W., E.M., J.G.P.J., K.V., L.M.O., A.M.K., A.O., C.G., A.B., R.R., M.S. (Mirjam Spengeler), F.R.S. and C.D.; supervision, C.D.; project administration, C.D.; funding acquisition, R.R. and C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Swiss National Science foundation, grant number 172911 to CD and 166119 to RR, and the APC was funded by the Swiss National Science foundation. RR and AB were funded by Art & Vie foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not experimental but part of clinical and pathological veterinary diagnostics. The animals were handled according to good ethical standards and Swiss legislation (Animal Welfare regulation: Tierschutzverordnung from 23 April 2008, last amended on 4 September 2018). The tissue for pathological examinations were collected after slaughtering. The aim was to identify the cause of the congenital disorder. All animals in this study were examined with the consent of their owners. Collection of blood samples was approved by the Cantonal Committee for Animal Experiments (Canton of Bern; permit 71/19).

Informed Consent Statement: Not applicable.

Data Availability Statement: The WGS data of the three sequenced cases can be found in the European Nucleotide Archive under the sample accession nos. SAMEA4644768 (case 2), SAMEA6528889 (case 3), and SAMEA6528891 (case 4).

Acknowledgments: The authors are grateful to Brauvieh Schweiz for assistance in sampling and organization of cattle for the clinical exams as well as for SNP array genotyping of all cases. The authors also wish to thank Swissgenetics for providing semen. Furthermore we thank Nathalie Besuchet and Simon Pot for expert assistance. We also acknowledge all bovine researchers contributing to the 1000 Bull Genomes Project and those who deposited cattle whole genome sequencing data into public databases.

Conflicts of Interest: Authors declare no conflict of interest.

References

1. Lavach, J.D. *Large Animal Ophthalmology*; Mosby: St. Louis, MO, USA, 1990.
2. Williams, D.L. Production animal ophthalmology. *Vet. Clin. N. Am. Food Anim. Pract.* **2010**, *26*, xi–xii. [CrossRef]
3. Williams, D.L. Welfare issues in farm animal ophthalmology. *Vet. Clin. N. Am. Food Anim. Pract.* **2010**, *26*, 427–435. [CrossRef] [PubMed]
4. Leipold, H.W. Congenital ocular defects in food-producing animals. *Vet. Clin. N. Am. Large Anim. Pract.* **1984**, *6*, 577–595. [CrossRef]
5. Gelatt, K.N.; Ben-Shlomo, G.; Gilger, B.C.; Hendrix, D.V.; Kern, T.J.; Plummer, C.E. *Veterinary Ophthalmology*, 6th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2021; pp. 1983–2054.
6. Williams, D.L. Congenital abnormalities in production animals. *Vet. Clin. N. Am. Food Anim. Pract.* **2010**, *26*, 477–486. [CrossRef]
7. He, X.; Li, Y.; Li, M.; Jia, G.; Dong, H.; Zhang, Y.; He, C.; Wang, C.; Deng, L.; Yang, Y. Hypovitaminosis A coupled to secondary bacterial infection in beef cattle. *BMC Vet. Res.* **2012**, *8*, 222. [CrossRef]
8. Murgiano, L.; Jagannathan, V.; Calderoni, V.; Joechler, M.; Gentile, A.; Drögemüller, C. Looking the cow in the eye: Deletion in the NID1 gene is associated with recessive inherited cataract in Romagnola cattle. *PLoS ONE* **2014**, *9*, e110628. [CrossRef] [PubMed]
9. Hollmann, A.K.; Dammann, I.; Wemheuer, W.M.; Wemheuer, W.E.; Chilla, A.; Tipold, A.; Schulz-Schaeffer, W.J.; Beck, J.; Schütz, E.; Brenig, B. Morgagnian cataract resulting from a naturally occurring nonsense mutation elucidates a role of CPAMD8 in mammalian lens development. *PLoS ONE* **2017**, *12*, e0180665.
10. Sun, W.; Li, S.; Xiao, X.; Wang, P.; Zhang, Q. Genotypes and phenotypes of genes associated with achromatopsia: A reference for clinical genetic testing. *Mol. Vis.* **2020**, *26*, 588–602. [PubMed]
11. Reicher, S.; Seroussi, E.; Gootwine, E. A mutation in gene CNGA3 is associated with day blindness in sheep. *Genomics* **2010**, *95*, 101–104. [CrossRef]
12. Ross, M.; Ofri, R.; Aizenberg, I.; Abu-Siam, M.; Pe'er, O.; Arad, D.; Rosov, A.; Gootwine, E.; Dvir, H.; Honig, H.; et al. Naturally-occurring myopia and loss of cone function in a sheep model of achromatopsia. *Sci. Rep.* **2020**, *10*, 19314. [CrossRef]
13. Komáromy, A.M. Day blind sheep and the importance of large animal disease models. *Vet. J.* **2010**, *185*, 241–242. [CrossRef] [PubMed]
14. Winkler, P.A.; Occelli, L.M.; Petersen-Jones, S.M. Large animal models of inherited retinal degenerations: A review. *Cells* **2020**, *9*, 882. [CrossRef] [PubMed]

15. Rothammer, S.; Seichter, D.; Förster, M.; Medugorac, I. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genom.* **2013**, *14*, 908. [CrossRef]
16. Signer-Hasler, H.; Burren, A.; Neuditschko, M.; Frischknecht, M.; Garrick, D.; Stricker, C.; Gredler, B.; Bapst, B.; Flury, C. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genet. Sel. Evol.* **2017**, *49*, 83. [CrossRef]
17. Bhati, M.; Kadri, N.K.; Crysnanto, D.; Pausch, H. Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. *BMC Genom.* **2020**, *21*, 27. [CrossRef]
18. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef]
19. Miyadera, K.; Acland, G.M.; Aguirre, G.D. Genetic and phenotypic variations of inherited retinal diseases in dogs: The power of within- and across-breed studies. *Mamm. Genome* **2012**, *23*, 40–61. [CrossRef] [PubMed]
20. Yeh, C.Y.; Goldstein, O.; Kukekova, A.V.; Holley, D.; Knollinger, A.M.; Huson, H.J.; Pearce-Kelling, S.E.; Acland, G.M.; Komáromy, A.M. Genomic deletion of CNGB3 is identical by descent in multiple canine breeds and causes achromatopsia. *BMC Genet.* **2013**, *14*, 27. [CrossRef]
21. Tanaka, N.; Delemotte, L.; Klein, M.L.; Komáromy, A.M.; Tanaka, J.C. A cyclic nucleotide-gated channel mutation associated with canine daylight blindness provides insight into a role for the S2 segment tri-Asp motif in channel biogenesis. *PLoS ONE* **2014**, *9*, e88768. [CrossRef]
22. Tanaka, N.; Dutrow, E.V.; Miyadera, K.; Delemotte, L.; MacDermaid, C.M.; Reinstein, S.L.; Crumley, W.R.; Dixon, C.J.; Casal, M.L.; Klein, M.L.; et al. Canine CNGA3 Gene Mutations Provide Novel Insights into Human Achromatopsia-Associated Channelopathies and Treatment. *PLoS ONE* **2015**, *10*, e0138943. [CrossRef]
23. Sidjanin, D.J.; Lowe, J.K.; McElwee, J.L.; Milne, B.S.; Phippen, T.M.; Sargan, D.R.; Aguirre, G.D.; Acland, G.M.; Ostrander, E.A. Canine *CNGB3* mutations establish cone degeneration as orthologous to the human achromatopsia locus ACHM3. *Hum. Mol. Genet.* **2002**, *11*, 1823–1833. [CrossRef]
24. Chen, X.T.; Huang, H.; Chen, Y.H.; Dong, L.J.; Li, X.R.; Zhang, X.M. Achromatopsia caused by novel missense mutations in the CNGA3 gene. *Int. J. Ophthalmol.* **2015**, *8*, 910–915.
25. Hagger, C. Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. *J. Anim. Breed. Genet.* **2005**, *122*, 405–413. [CrossRef] [PubMed]
26. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]
27. Sargolzaei, M.; Chesnais, J.P.; Schenkel, F.S. FImpute—An efficient imputa- tion algorithm for dairy cattle populations. *J. Dairy Sci.* **2011**, *94*, 421.
28. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
29. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef]
30. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]

Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, visualizations, writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.12923
John Wiley and Sons License Number:	5236040564250

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

ORIGINAL ARTICLE

doi: 10.1111/age.12923

Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome

I. M. Häfliger*, N. Wiedemar*, T. Švara†, J. Starič†, V. Cociancich†, K. Šest†, M. Gombač†, T. Paller†, J. S. Agerholm‡ and C. Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland. †Veterinary Faculty, University of Ljubljana, Ljubljana 1000, Slovenia. ‡Department of Veterinary Clinical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg C 1870, Denmark.

Summary

The pulmonary hypoplasia and anasarca syndrome (PHA) is a congenital lethal disorder, which until now has been reported in cattle and sheep. PHA is characterized by extensive subcutaneous fetal edema combined with hypoplasia or aplasia of the lungs and dysplasia of the lymphatic system. PHA is assumed to be of genetic etiology. This study presents the occurrence of PHA in two different cattle breeds and their genetic causation. Two PHA cases from one sire were observed in Slovenian Cika cattle. Under the assumption of monogenic inheritance, genome-wide homozygosity mapping scaled down the critical regions to 3% of the bovine genome including a 43.6 Mb-sized segment on chromosome 6. Whole-genome sequencing of one case, variant filtering against controls and genotyping of a larger cohort of Cika cattle led to the detection of a likely pathogenic protein-changing variant perfectly associated with the disease: a missense variant on chromosome 6 in *ADAMTS3* (NM_001192797.1: c.1222C>T), which affects an evolutionary conserved residue (NP_001179726.1: p.(His408Tyr)). A single PHA case was found in Danish Holstein cattle and was whole-genome sequenced along with its parents. However, as there was no plausible private protein-changing variant, mining for structural variation revealed a likely pathogenic trisomy of the entire chromosome 20. The identified *ADAMTS3* associated missense variant and the trisomy 20 are two different genetic causes, which shows a compelling genetic heterogeneity for bovine PHA.

Keywords cattle, chromosomal aberration, genetic disorder, Mendelian, monogenic, precision medicine, trisomy, rare disease

Introduction

Pulmonary hypoplasia with anasarca syndrome (PHA) is a rare congenital lethal syndrome described in cattle (OMIA 001562-9913) and sheep (OMIA 000493-9940). Known cattle breeds with reported cases of PHA are Belted Galloway, Dexter, Maine-Anjou, Shorthorn and Hereford cattle (Windsor *et al.* 2006; Whitlock *et al.* 2008; Agerholm & Arnbjerg 2011). The affected fetuses show diffuse severe subcutaneous edema, accumulation of fluids in the body cavities (anasarca) and a dysplasia of the lymphatic system in addition to pulmonary hypoplasia or aplasia. Placental edema might be present (Windsor *et al.* 2006; Agerholm & Arnbjerg 2011). The described Galloway case additionally showed palatoschisis, cardiac malformations and cryptorchidism (Agerholm & Arnbjerg 2011). PHA is associated with a high percentage of abortion, but some cases are delivered near term (Whitlock & Coffman 2015). The extensive subcutaneous edema causes an increased and abnormal body size that often prevents the fetus from entering the birth canal properly. If the calving is not recognized and proper assistance delivered, the fetus dies and undergoes postmortem degradation, which causes a life-threatening condition for the dam (Windsor *et al.* 2006; Agerholm & Arnbjerg 2011). PHA is therefore associated with an increased maternal mortality rate, especially because it has mainly been reported in beef cattle that are usually monitored less frequently than dairy cattle.

In Dexter cattle all of the described cases in the publication of Windsor *et al.* (2006) derived from one

Address for correspondence

C. Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland.
E-mail: cord.droegemueller@vetsuisse.unibe.ch

Accepted for publication 24 January 2020

cow, which is assumed to be the founder. Therefore, a recessive mode of inheritance was suspected. In the Galloway case, pedigree analysis identified a common maternal and paternal ancestor, which could also indicate a recessive mode of inheritance (Agerholm & Arnbjerg 2011), but the fetus was too degraded to allow analysis of its DNA. Furthermore, PHA cases have been observed in Maine-Anjou and Shorthorn cattle and a single case in Hereford (Whitlock *et al.* 2008). A DNA test for the PHA forms found in Dexter, Shorthorn and Maine-Anjou cattle are offered commercially, although the causative variant has not been published so far (https://genetics.zoetis.com/canada/products/beef/pulmonary-hypoplasia-with-anasarca.aspx; https://www.neogen.com/en/geneseek-announces-next-generation-of-dna-technology-geneseek-genomic-profiler-bovine-hd).

Moreover, PHA has been described as hydrops foetalis in sheep (OMIA 000493-9940; Monteagudo *et al.* 2002; Alleaume *et al.* 2012). The cases reported by Monteagudo *et al.* (2002) derived from one ram which had been selected from the same flock, indicating a recessive inheritance. Similar to the bovine cases the ovine fetuses were oversized and showed anasarca and hypoplastic lungs (Monteagudo *et al.* 2002; Alleaume *et al.* 2012). Skeletal muscle hypoplasia and brachygnathia were found in some cases (Monteagudo *et al.* 2002).

In the present study we describe the genetic cause of independent cases of PHA in two breeds: Cika cattle, an autochthonous Slovenian breed and with PHA cases previously described by Švara *et al.* (2016), and Holstein cattle from Denmark.

Materials and methods

Animals

EDTA-blood samples of two Slovenian Cika PHA-affected fetuses, their sire and one dam, as well as seven closely related controls, were available. In addition, a further 158 DNA samples were taken from the biobank of the Veterinary Faculty of the University of Ljubljana as controls. The morphology and familial relationship (Fig. 2a) of the two cases have previously been reported by Švara *et al.* (2016). Whereas both cases and the available parents had been sent for SNP genotyping, only the case where both parents were genotyped was whole-genome sequenced (Fig. 2a). From a Danish Holstein PHA case, a fetal spleen sample, EDTA-blood of the dam and semen of the sire were available and the trio was whole-genome sequenced. DNA was isolated either from EDTA-blood using the Nucleon Bacc2 kit (GE Healthcare) or from semen, hair roots or tissue samples using QIAGEN's DNeasy kit according to the manufacturer's instructions and using the Maxwell RSC instrument (Promega).

All three PHA-affected cases were tested homozygous wt for the variants of the commercial DNA test for bovine PHA.

Homozygosity mapping

High-density SNP genotyping was carried out for the two affected Cika fetuses and two obligate carrier animals on the Illumina BovineHD BeadChip array including 777 962 SNP markers at GeneSeek (https://genomics.neogen.com/en/illumina-bovine-hd-beadchip). Quality control was performed based on the assumptions that the genotype quality per animals is good, at least one of the genotyped individuals is heterozygous and all animals were genotyped at each particular SNP. Even though no sample was excluded, a total of 354 381 variants were removed owing to there being no heterozygous animal in the dataset and 16 313 variants were removed owing to missing genotypes. With a total of 389 586 remaining markers, homozygosity mapping was performed for the two affected offspring using the software PLINK v1.9 (Purcell *et al.* 2007) with the commands *--homozyg-kb* 100 (considering homozygous segments of at least 100 kb), *--homozyg-match* 0.95 (for allelic matching between both cases) and *--homozyg-group* (for generating an overlap-file), resulting in shared runs of homozygosity (ROH).

Whole-genome sequencing

We prepared individual PCR-free fragment libraries with approximately 400 bp inserts which were sequenced for 150 bp paired-end reads. A single Cika PHA case was then sequenced on the Illumina HiSeq2500 at 8.4 read depth and the whole Holstein family trio was sequenced on the Illumina NovaSeq6000 at a coverage of approximately 20×. The WGS data is available under the study accession no. PRJEB18113 at the European Nucleotide Archive [www.ebi.ac.uk/ena; sample accessions SAMEA32989918 (Cika case), SAMEA5160152 (Holstein case), SAMEA5159839 (dam), SAMEA5159890 (sire)]. We mapped the reads to the latest cattle reference genome ARS-UCD1.2 (https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1) and called for single nucleotide variants (SNVs) and small insertions and deletions. The pipeline used to process the fastq files to binary alignment map and to genomic variant call format (GVCF) files was the proposed workflow provided by the 1000 Bull Genomes Project (run 7; www.1000bullgenomes.com; Hayes & Daetwyler 2019). The resulting individual GVCF files were merged into one large variant call format (VCF) file using CombineGVCFs and CatVariants of GATK v3.8 (DePristo *et al.* 2011). SNVs and insertions and deletions were called using GenotypeGVCF of GATK version 3.8 (DePristo *et al.* 2011) and were given a quality label based on the best practice recommendations in GATK using the VariantFiltration of GATK v3.8 (DePristo *et al.* 2011). Finally, their impact was functionally annotated with SNPEFF v4.3 (Cingolani *et al.* 2012), using the NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_e

uk/Bos_taurus/106/), which resulted in the final VCF file, including all individual variants and their functional annotations.

WGS data analysis

In addition to the four animals sequenced for this study, there were another 395 control genomes available (Table S1). The control genomes consist of animals from 23 different cattle breeds and nine cross-bred animals. The final VCF file was filtered genome-wide based on the assumed mode of inheritance (breed dependent) using a case–control approach. The investigation for large structural variants was performed using SAMTOOLS v0.1.19-44428cd (Li *et al.* 2009) in order to calculate the read depth per base. Further, R (http://www.R-project.org/) was used to calculate the average read depth over a window of 250 kb and the *manhattan* function of the QQMAN package (Turner 2014) to visualize the read depth over the whole genome in one plot. Integrative Genomic Viewer software (Robinson *et al.* 2011) was used for visual inspection of sequence variants. The read depth data of the case was further evaluated statistically within an ANOVA analysis using the MIXLM v1.2.3 package in R (https://CRAN.R-project.org/package=mixlm). Thereby, we investigated whether there is any variation in read depth between the chromosomes. Using the same package, a *post-hoc* Tukey test was performed to check which chromosomes differ significantly from each other and the results were visualized using the *cld* function (https://CRAN.R-project.org/package=mixlm).

In order to evaluate the parentage of the sequenced trio, the SNVs were extracted based on a MAF of 0.01, the PASS quality prediction of GATK, and stored in PLINK format using VCFTOOL version 0.1.15 (Danecek *et al.* 2011). PLINK version 1.9 (Purcell *et al.* 2007) was used to perform further quality control by limiting the amount of missing calls per animal and per variant to a maximum of 10 and 5% respectively. With the function *--check-sex* the sex of the individuals was compared, with *--genome* the paternity/maternity was tested, and with *--mendel* the variants not following a Mendelian inheritance pattern were detected.

Genotyping of candidate variants

Primers (Table S3) were designed with PRIMER3 software (Untergasser *et al.* 2012) after masking of repetitive sequences with RepeatMasker (http://www.repeatmasker.org). Sanger sequencing of the PCR products was performed using the BIGDYE TERMINATOR version 3.1 cycle sequencing kit (LifeTechnologies) after purification with rAPid alkaline phosphatase (Roche) and exonuclease I (New England BioLabs). Sequencing products were resolved on an ABI 3730 capillary sequencer (LifeTechnologies) and the obtained sequence data were analyzed with SEQUENCHER 5.1 software.

Results

PHA in Cika cattle

The morphology of the Cika cases has been reported previously (Švara *et al.* 2016). Briefly, the lesions were typical for PHA with severe diffuse subcutaneous edema with multiple fluid-filled cysts, accumulation of fluid in the body cavities, hypoplastic lungs and dysplasia of the lymphatic tissues (Fig. 1). Pedigree analysis (Fig. 2a) revealed a highly inbred family, where the cases resulted from a mother–son and an aunt–nephew mating, indicating monogenic recessive inheritance.

Homozygosity mapping revealed a total of 18 ROH segments, which are likely to harbor the recessive mutation (Fig. 2b). In these candidate regions both cases share the same allelic group and cover a combined 90.8 Mb, corresponding to roughly 3% of the bovine genome (Table S2). Subsequently WGS of one of the cases was performed. Owing to the assumed recessive inheritance and the lethal effect of the mutation, we hypothesized that most likely a protein-changing variant affecting the coding sequence of a gene would be responsible for PHA in Cika cattle. The VCF file was screened for homozygous private variants in the case. A total of 10 130 homozygous variants that passed the quality filtering were detected in the case. After filtering variants against the 398 control genomes (including the Holstein PHA family) and restricting the oness to be located within the ROH candidate regions, 10 were left. These variants were further checked for their occurrence in a global control cohort of 3103 genomes of other breeds (1000 Bull Genomes Project run 7; www.1000bullgenomes.com), which revealed four protein-changing variants only present in the Cika PHA-affected fetus (Table S3).

These candidate variants were genotyped by Sanger sequencing. This led to the further exclusion of three variants, which did not segregate perfectly with the phenotype, as non-affected Cika controls were homozygous for the variants (Table S3). Finally, one single candidate variant was left: a G>A variant on chromosome 6 at position 87 462 016 located in the *ADAMTS3* gene. For this variant both cases were homozygous AA and seven out of 169 Cika cattle, including the two obligate carriers, were heterozygous for this variant. As the gene is on the reverse strand of the chromosome, this leads to a C>T exchange on a transcript level (NM_001192797.1: c.1222C>T; Fig. 3a). The nucleotide exchange is non-synonymous and results in an amino acid exchange of histidine to tyrosine at position 408 of the protein (NP_001179726.1: p.(His408Tyr); Fig. 3a). The affected amino acid is highly conserved among different species (Fig. 3b) and it is predicted to be probably damaging by PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) and not tolerated by SIFT (https://sift.bii.a-star.edu.sg/). Multiple sequence alignment across different species shows a high degree of conservation of the concerned

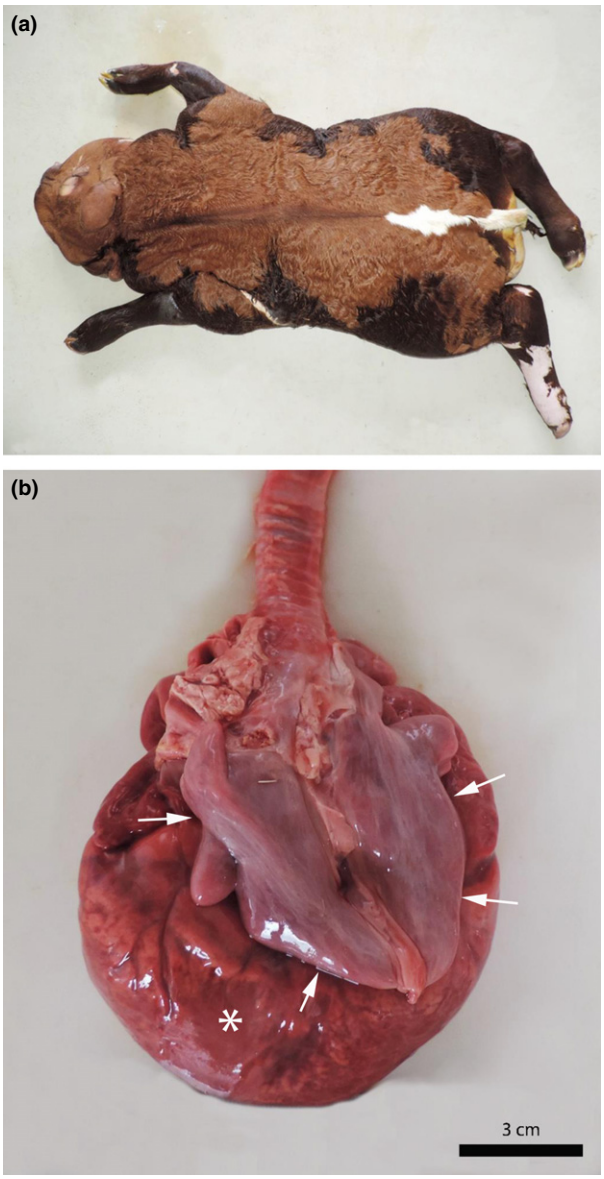


Figure 1 Gross morphology of pulmonary hypoplasia and anasarca syndrome (PHA) in a Cika fetus. (a) The body is severely deformed owing to extensive subcutaneous edema. (b) Specimen showing the heart and lungs. Whereas the heart (asterix) is enlarged and rounded owing to marked dilatation of the heart chambers, the lungs (arrows) are extremely reduced in size. Bar = 3 cm.

amino acid in PHA-affected Cika cattle (Fig. 3b) and the altered residue is situated in a functionally important metalloproteinase domain (Fig. 3c).

PHA in Holstein cattle

The female Holstein fetus was aborted at gestation day 235. The pregnancy was the result of insemination with semen of a US Holstein sire and the parents were not related within at least three generations. This was the only PHA case related to this sire that was reported to the surveillance system for congenital bovine disorders in Denmark (Agerholm *et al.* 1993). Owing to summer heat, the fetus was frozen at −20°C until necropsy was possible.

The fetus, weighing 17.7 kg, was examined after thawing. It was deformed owing to diffuse severe subcutaneous edema, which caused large localized soft tissue swellings, especially in the head and dorsal part of the body (Fig. 4a). Upon incision these swellings consisted of either edematous connective tissue with multiple small cysts or a large solitary cyst (Fig. 4b). All pulmonary lobes were developed but the lung size was significantly reduced. The heart had a high interventricular septal defect, dilation of both ventricles and hypertrophy of the left myocardium. A palatoschisis was present. The amount of fluid in the body cavities was considered to be within normal range.

To evaluate the genetic etiology, WGS of the case along with the sire and dam was performed. The variant data confirmed the parentage of the fetus. Under the assumption of recessive inheritance, the VCF file was screened for protein-changing variants in the trio, leaving 1042 variants passing the GATK quality filter. After filtering against 396 control genomes (including the Cika PHA case), 13 variants were left for which no homozygous genotype was detected. When comparing these remaining variants with the 1000 Bull Genome Project data, finally a single missense variant in the *SERPINB6* gene remained. This variant never occurred in homozygous state in 3102 control genomes, but was heterozygous present in 51 animals of nine different breeds. The *SERPINB6*-associated base exchange from guanine to thymine (NM_174789.1: c.1042G>T) leads to an amino acid exchange from valine to leucine (XM_024983399.1: p.(Val348Leu)). This *SERPINB6* variant was of uncertain significance, and therefore, a *de novo* pathogenic mutation was hypothesized.

Applying this hypothesis, the PHA-affected Holstein case was filtered for protein-changing variants in heterozygous state and compared with its parents, which are assumed to be non-carriers of these variants. This led to the detection of 744 putative *de novo* variants. After comparison against 396 control genomes, nine protein-changing candidate variants remained, and using the 1000 Bull Genomes Project data for validation, additional two variants could be excluded. Visual inspection of these seven variants indicated that apparently six represent false positives. The single remaining variant was a missense variant in the *RAB9A* gene (NM_001075538.2: c.565C>T), leading to an amino acid substitution from arginine to cysteine (NP_001069006.1: p.(Arg189Cys)). Owing to the functional association of the *RAB9A* gene with pigmentation phenotypes (Mahanty *et al.* 2016), this variant of uncertain significance was not considered as pathogenic for PHA.

However, visual inspection for large structural variants in the genome was performed after plotting the average read depth across the whole genome for the Holstein trio (Fig. 5). As proof of concept, this method confirms the sex of the

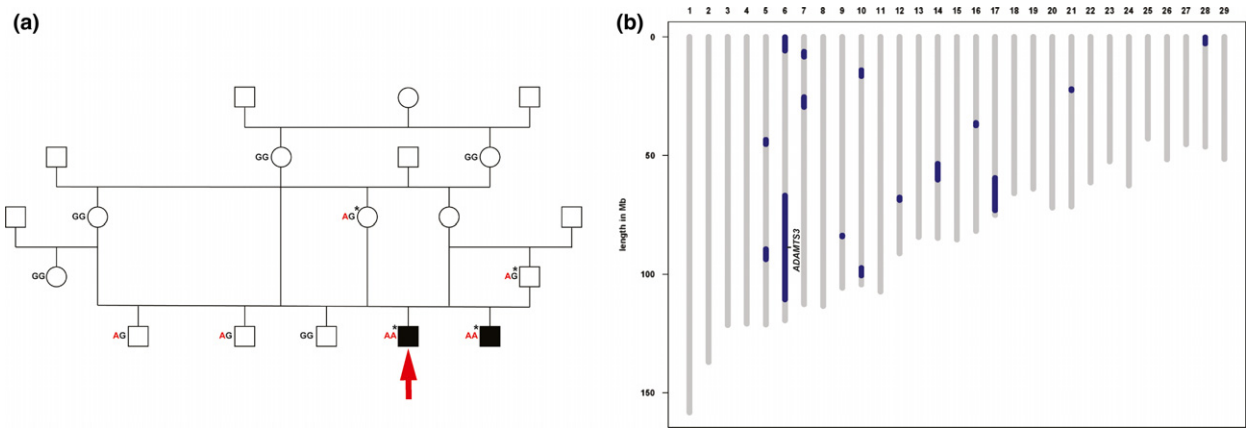


Figure 2 Pedigree and genome-wide homozygosity mapping of the affected Cika cattle. (a) The two PHA-affected fetuses (filled symbols) are the results of a mother–son and an aunt–nephew mating. The animals shown with an asterisk were genotyped with the SNP array and the genome of an affected fetus was sequenced (indicated with an arrow). The individual genotypes of the *ADAMTS3* variant are shown, whereby the variant allele A is indicated in red and the wt allele G in black. (b) Homozygous regions shared by both cases are displayed in blue. Note the by far largest homozygous segment of 43.6 Mb on chromosome 6, *ADAMTS3*, is located within this segment.

animals as the sire shows a decreased read depth of the X chromosome. Furthermore, we detected a trisomy encompassing the entire chromosome 20 (Fig. 5). The ANOVA results showed that there was a significant difference in read depth from the mean of at least one chromosome (F-value = 93.859, DF = 28, P-value = <2.2e − 16). The pairwise Tukey test demonstrated that each chromosome differed significantly from some other chromosomes, but only chromosome 20 differed from all other chromosomes (Table S4).

Discussion

In ruminants, PHA has been described in sheep and in cattle. The described cases were aborted late in gestation or delivered at term. PHA cases are described as showing extensive edema and an increased body size in sheep (Monteagudo *et al.* 2002; Alleaume *et al.* 2012) and in cattle (Windsor *et al.* 2006; Whitlock *et al.* 2008; Agerholm & Arnbjerg 2011). In all cases, pathology revealed undersized lungs and dysplasia in the lymphatic system. Interestingly, even though the genetic origin is mostly assumed to be of recessive inheritance, there are cases in cross-bred sheep where a recessive inheritance would be a great coincidence. Furthermore, so far no causative variant has been published for PHA either in sheep or in cattle. The absence of the unreported but commercially tested variants for bovine PHA indicates that the PHA-affected calves presented here are unrelated to those that led to the commercial gene test development.

Herein with the missense variant in *ADAMTS3*, we present for the first time molecular evidence for the assumed simple Mendelian inheritance of bovine PHA in Cika cattle, a local breed of Slovenia. The *ADAMTS3* gene belongs to the *ADAMTS* proteases, a gene family which is characterized by the presence of a disintegrin and metalloproteinase (ADAM) domain and thrombospondin 1 repeats (Tang & Hong 1999). In previous reports, ADAMTS3 was proposed to be the major procollagen II-processing enzyme, and a possible role in musculoskeletal development was discussed (Le Goff *et al.* 2006; Marouli *et al.* 2017). In mice *Adamts3* transcripts are predominantly expressed in cartilage throughout murine chondrogenesis and bone ossification (Dubail & Apte 2015). In humans and dogs, *ADAMTS3* is expressed in the cartilage of the upper airway (Jeltsch *et al.* 2014; Marchant *et al.* 2019). *ADAMTS3* KO mice have been reported, even though their phenotypes differ substantially, perhaps owing to differences in design and/or genetic background; in both studies abnormalities in lymphatic development and subcutaneous edema were reported (Janssen *et al.* 2016; Ogino *et al.* 2017). Whereas mice described by Janssen *et al.* (2016) died predominantly as embryos, the mice described by Ogino *et al.* (2017), tended to die perinatally owing to apparent breathing problems or present abnormal rib development and significantly shortened skulls. In a human family, two variants of *ADAMTS3* have been shown to cause autosomal recessive-inherited Hennekam lymphangiectasia–lymphedema syndrome 3, a condition characterized by lymphedema and distinct facial features, including hypertelorism and a flat nasal bridge (OMIM 618154; Hennekam *et al.* 1989; Brouillard *et al.* 2017). These findings support the *ADAMTS3* missense variant as the most likely pathogenic variant. Previous reports already suspected the primary cause of the PHA phenotype to be associated with dysplasia of the lymphatic system (Whitlock *et al.* 2008; Alleaume *et al.* 2012). The presented features of the PHA phenotype in Cika cattle (Švara *et al.* 2016) support this assumption and the identified *ADAMTS3* variant confirms its origin to a disruption of the lymphangiogenesis. It can be speculated that the described PHA-associated *ADAMTS3* variant leads to a complete loss of function.

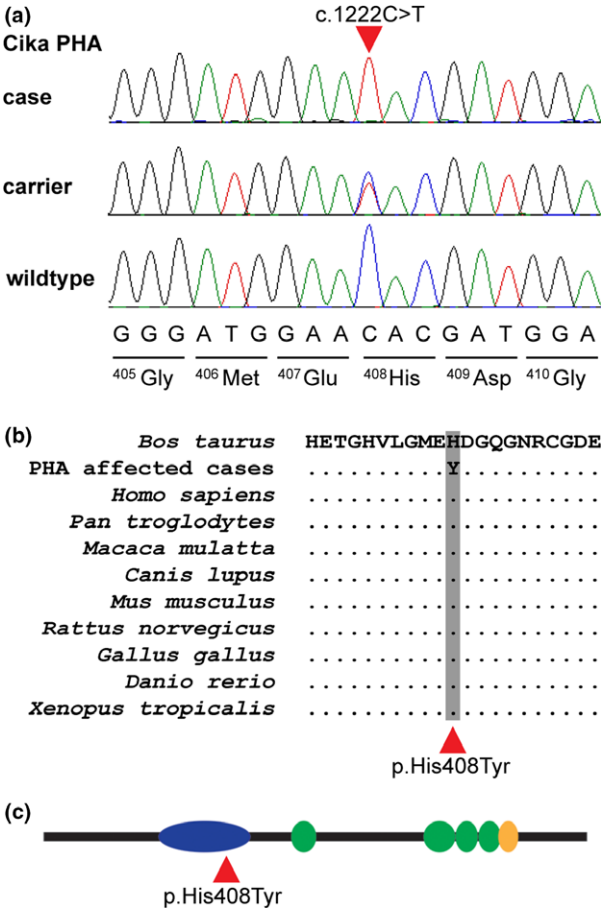


Figure 3 PHA-associated *ADAMTS3* missense variant in Cika cattle. (a) Genotypes identified via Sanger sequencing showing the different genotypes. (b) Multiple protein sequence alignment across species. (c) Schematic representation of the ADAMTS3 protein according to Kelwick et al. (2015). The affected metalloproteinase domain is shown in blue, the thrombospondin type 1 motifs in green and the protease and lacunin module in yellow.

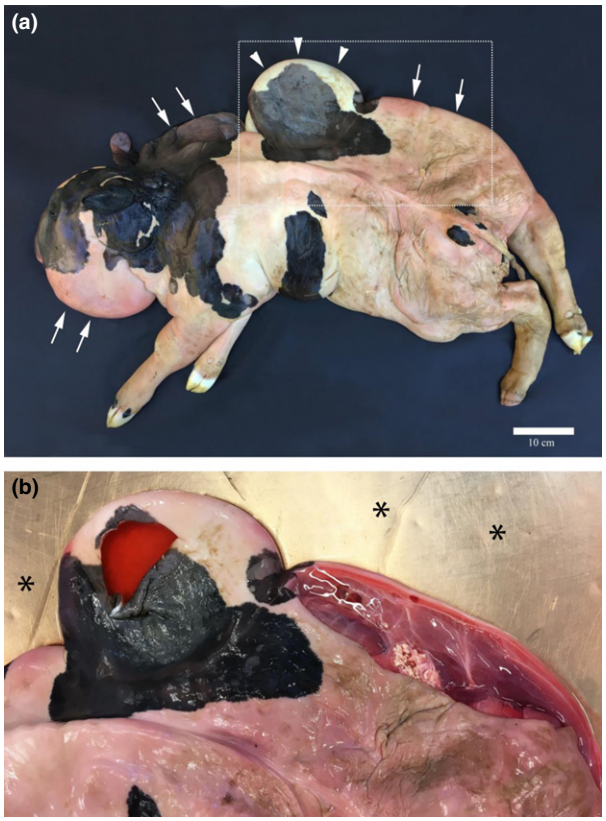


Figure 4 Gross morphology of a case of PHA in a Holstein fetus. (a) The body is deformed, especially the head and dorsal part of the body, owing to subcutaneous swellings (arrows) and a subcutaneous cyst (arrow heads). (b) Cut section through lesions in the dorsal part of the fetal body as indicated in (a). The cyst has been opened and serous fluid is oozing from the cavity (asterisk); the swelling to the right consists of severe subcutaneous edema.

By genotyping we were able to exclude the two maternal grandmothers as the founder of the Cika PHA variant (Fig. 2). Therefore, we suppose the founder of the Cika PHA mutation is the common maternal grandsire and the age of the variant only being two generations. The by far largest shared ROH segment is the 43 Mb region on chromosome 6, harboring the identified *ADAMTS3* variant, which also supports a recent mutation event. The occurrence of PHA in the cohort was due to the highly inbred family structure. Genotyping for the described variant in Cika cattle and selective breeding will help to reduce the allele frequency in the population and minimize the risk of more PHA cases.

The presented individual case of PHA observed in Holstein cattle is most likely not caused by a recessive mutation as we detected a chromosomal aberration. Formally we cannot exclude the two protein-changing variants in *RAB9A* and *SERPINB6* as the causative variant in the Holstein case. However, functionally *RAB9A* is involved in protein transport between the endosomes to the melanosomes and its KO had been shown to lead to hypopigmentation *in vitro* (Mahanty *et al.* 2016). On the other hand, the variant in *SERPINB6* does not seem to be a true *de novo* variant, as it appears in more than 50 animals across several breeds. Furthermore, *SERPINB6* is associated in human with hearing impairment (Sirmaci *et al.* 2010). Based on this information, we assume these variants to be probably not causal for the PHA case in the Holstein breed.

Aneuploidies, defined as the copy number variation of whole chromosomes, lead to changes in gene expression. Chunduri & Storchova (2019) described the effects of a trisomic chromosome in raising the gene expression and disturbing protein homeostasis in the cell and finally impairing the functioning of the cell cycle. Thus, it is more plausible that the extra copy of chromosome 20 is causing the disorder in this isolated case. No further cases of PHA have been reported in Holstein cattle in Denmark associated with the sire of the affected fetus, which supports this conclusion. Even though cases of trisomy have been mostly

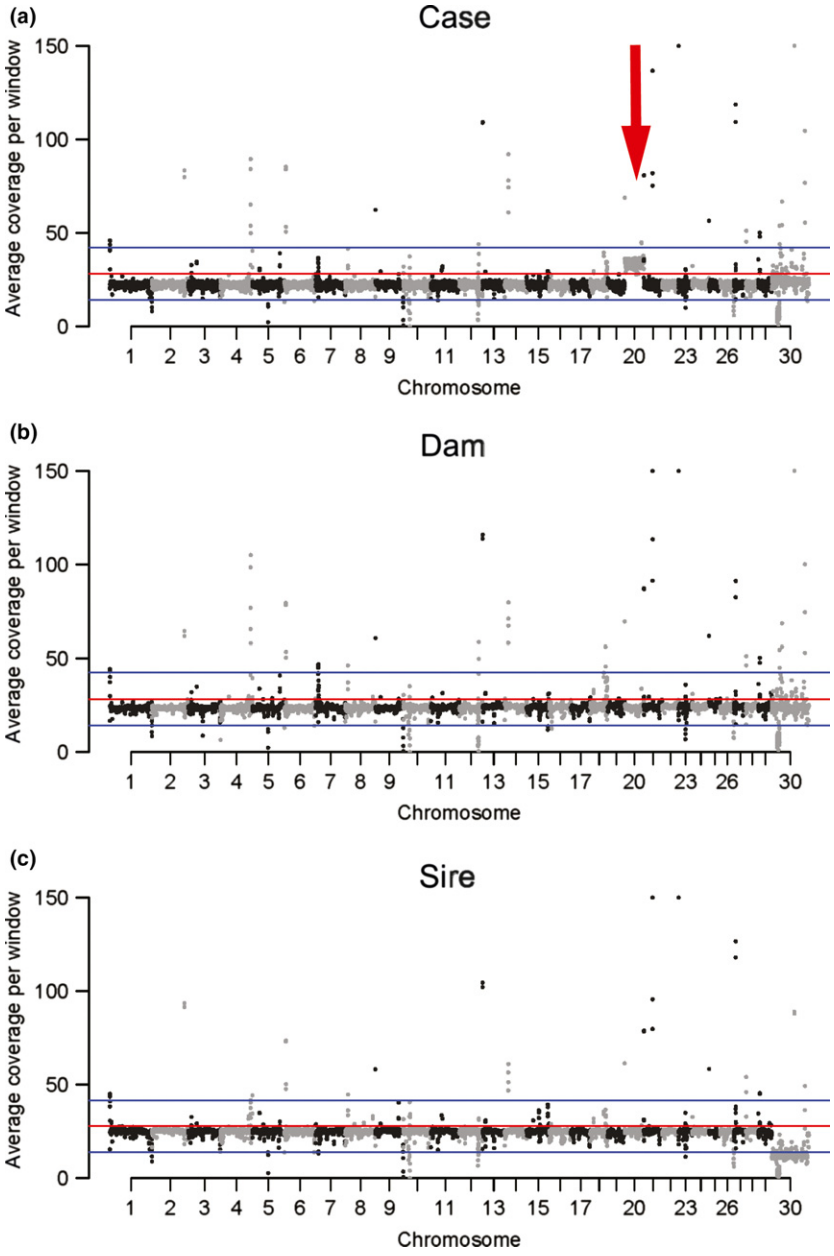


Figure 5 PHA-associated trisomy of chromosome 20 in Holstein cattle based on WGS data. Average read depth of 250 kb windows plotted for the affected fetus (a), the dam (b) and the sire (c). The red arrow indicates the increased sequence coverage across the whole chromosome 20.

described in humans, there are few cases in cattle described in the literature for cattle chromosomes 18, 20, 21, 22 and 28 (Herzog *et al.* 1977; Mayr *et al.* 1985; Agerholm & Christensen 1993; Schmutz *et al.* 1996; Gallagher *et al.* 1999; Iannuzzi *et al.* 2001). All of these studies used traditional cytogenetic approaches to describe the chromosomal disorder. Unfortunately, in our case there was no suitable material stored for cell culture. The descriptions of all of the reported trisomies had phenotypic signs in common such as facial dysplasia, brachygnathia and/or palatoschisis, and malformations of the limbs (Herzog *et al.* 1977; Schmutz *et al.* 1996; Iannuzzi *et al.* 2001).

Autosomal trisomy 20 in cattle has been reported previously by Gallagher *et al.* (1999), who by coincidence found such a case in the fetus of a slaughtered pregnant beef cow. The fetus displayed severe anasarca and cranio-facial malformation including cranioschisis, palatoschisis and cheiloschisis. The present Holstein fetus with trisomy 20 also had anasarca, although the head was less malformed as merely palatoschisis was present. Malformation of the heart and lung was not mentioned in the report by Gallagher *et al.* (1999). The findings in these two cases show that the phenotype of trisomy 20 in cattle is externally characterized by severe anasarca causing an abnormally shaped body, but various concomitant lesions may also be present. Therefore, trisomy 20 should be suspected in PHA cases. For its detection, karyotyping should be attempted as long as cytogenetics is economic and time efficient.

Alternatively, it has been shown recently that array genotyping can be used for the detection of large copy number variants through the measures of the B-allele frequency and the logged ratio of observed probe intensity to expected intensity (Grahofer *et al.* 2019).

The comparative map of cattle and human chromosomes identifies two conserved regions between human chromosome 5 and bovine chromosome 20 (Band *et al.* 2000). More precisely, the whole short arm until 74 Mb and the end of the long arm (168.1–173.3 Mb) from human chromosome 5 correspond to the bovine chromosome 20. As this human chromosome is much larger than the corresponding bovine chromosome, aneuploidies cannot be compared directly. However, a review of partial trisomy of the short arm of human chromosome 5 shows its association with mental as well as growth retardation (Cervera *et al.* 2005).

In conclusion, PHA in sheep and cattle is a previously described genetic disorder; however this study is the first to report its genetic background. The described cases in two different breeds were independently diagnosed with the same disorder, but the genetic analysis revealed the cause to be of different origin. Apparently rare cases of congenital disorders resembling monogenic Mendelian diseases might also be caused by chromosomal abnormalities.

Acknowledgements

The authors are grateful to the farmers for reporting the cases and providing samples for this sutdy. We thank Nathalie Besuchet-Schmutz and Jan Henkel for expert technical assistance. We acknowledge the assistance of the Next Generation Sequencing Platform of the University of Bern and the Interfaculty Bioinformatics Unit of the University of Bern for performing the whole-genome re-sequencing experiments and providing high-performance computing infrastructure respectively. This study was funded by the Danish Bovine Genetic Disease Programme and the Swiss National Science Foundation. We thank Jonathan E. Beever (University of Illinois) for confirming that the herein described variants differ from the unreported but commercially tested variants for bovine PHA in other breeds.

References

Agerholm J.S., Arnbjerg J. (2011) Pulmonary hypoplasia and anasarca syndrome in a belted Galloway calf. *Veterinary Record* **168**, 190.

Agerholm J.S., Christensen K. (1993) Trisomy-22 in a calf. *Zentralblatt für Veterinärmedizin Reihe A* **40**, 576–81.

Agerholm J.S., Basse A., Christensen K. (1993) Investigations on the occurrence of hereditary diseases in the Danish cattle population 1989–1991. *Acta Veterinaria Scandinavica* **34**, 245–53.

Alleaume C., Strugnell B., Spooner R., Schock A. (2012) Hydrops foetalis with pulmonary hypoplasia in Cheviot and Cheviot-Texel cross-lambs. *Veterinary Record* **171**, 646.

Band M.R., Larson J.H., Rebeiz M. *et al.* (2000) An ordered comparative map of the cattle and human genomes. *Genome Research* **10**, 1359–68.

Brouillard P., Dupont L., Helaers R., Coulie R., Tiller G.E., Peeden J., Colige A., Vikkula M. (2017) Loss of ADAMTS3 activity causes Hennekam lymphangiectasia-lymphedema syndrome 3. *Human Molecular Genetics* **26**, 4095–104.

Cervera M., Sanchez S., Molina B., Alcantara M.A., Del Castillo V., Carnevale A., Angel A.G.D. (2005) Trisomy of the short arm of chromosome 5 due to a de novo inversion and duplication (5) (p15.3 p13.3). *American Journal of Medical Genetics Part A* **136**, 381–5.

Chunduri N.K., Storchova Z. (2019) The diverse consequences of aneuploidy. *Nature Cell Biology* **21**, 54–62.

Cingolani P., Platts A., Wang L.L., Coon M., Nguyen T., Wang L., Land S.J., Lu X., Ruden D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92.

Danecek P., Auton A., Abecasis G. *et al.* (2011) The variant call format and VCFtools. *Bioinformatics* **27**, 2156–8.

DePristo M.A., Banks E., Poplin R. *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491–8.

Dubail J., Apte S.S. (2015) Insights on ADAMTS proteases and ADAMTS-like proteins from mammalian genetics. *Matrix Biology* **44**, 24–37.

Gallagher D.S., Lewis B.C., De Donato M., Davis S.K., Taylor J.F., Edwards J.F. (1999) Autosomal trisomy 20 (61, XX,+20) in a malformed bovine fetus. *Veterinary Pathology* **36**, 448–51.

Le Goff C., Somerville R.P., Kesteloot F., Powell K., Birk D.E., Colige A.C., Apte S.S. (2006) Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis. *Development* **133**, 1587–96.

Grahofer A., Letko A., Häfliger I.M., Jagannathan V., Ducos A., Richard O., Peter V., Nathues H., Drögemüller C. (2019) Chromosomal imbalance in pigs showing a syndromic form of cleft palate. *BMC Genomics* **20**, 349.

Hayes B.J., Daetwyler H.D. (2019) 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual Review of Animal Biosciences* **7**, 89–102.

Hennekam R.C.M., Geerdink R.A., Hamel B.C.J., Hennekam F.A.M., Kraus P., Rammeloo J.A., Tillemans A.A.W. (1989) Autosomal recessive intestinal lymphangiectasia and lymphedema, with facial anomalies and mental-retardation. *American Journal of Medical Genetics* **34**, 593–600.

Herzog A., Hohn H., Rieck G. (1977) Survey of recent situation of chromosome pathology in different breeds of German cattle. *Genetics Selection Evolution* **9**, 471–91.

Iannuzzi L., Di Meo G.P., Leifsson P.S., Eggen A., Christensen K. (2001) A case of trisomy 28 in cattle revealed by both banding and FISH-mapping techniques. *Hereditas* **134**, 147–51.

Janssen L., Dupont L., Bekhouche M. *et al.* (2016) ADAMTS3 activity is mandatory for embryonic lymphangiogenesis and regulates placental angiogenesis. *Angiogenesis* **19**, 53–65.

Jeltsch M., Jha S.K., Tvorogov D., Anisimov A., Leppanen V.M., Holopainen T., Kivela R., Ortega S., Karpanen T., Alitalo K. (2014) CCBE1 enhances lymphangiogenesis via A disintegrin

and metalloprotease with thrombospondin motifs-3-mediated vascular endothelial growth factor-C activation. *Circulation* **129**, 1962–71.

Kelwick R., Desanlis I., Wheeler G.N., Edwards D.R. (2015) The ADAMTS (A disintegrin and metalloproteinase with thrombospondin motifs) family. *Genome Biology* **16**, 113.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–9.

Mahanty S., Ravichandran K., Chitirala P., Prabha J., Jani R.A., Setty S.R.G. (2016) Rab9A is required for delivery of cargo from recycling endosomes to melanosomes. *Pigment Cell & Melanoma Research* **29**, 43–59.

Marchant T.W., Dietschi E., Rytz U. *et al.* (2019) An ADAMTS3 missense variant is associated with Norwich Terrier upper airway syndrome. *PLoS Genetics* **15**, e1008102.

Marouli E., Graff M., Medina-Gomez C. *et al.* (2017) Rare and low-frequency coding variants alter human adult height. *Nature* **542**, 186–90.

Mayr B., Krutzler H., Auer H., Schleger W., Sasshofer K., Glawischnig E. (1985) A viable calf with trisomy-22. *Cytogenetics and Cell Genetics* **39**, 77–9.

Monteagudo L., Lujan L., Tejedor T., Climent S., Acin C., Navarro A., Arrugaa M.V. (2002) Fetal anasarca (Hydrops foetalis) associated with lymphoid tissue agenesis possibly due to an autosomal recessive gene defect in sheep. *Theriogenology* **58**, 1219–28.

Ogino H., Hisanaga A., Kohno T. *et al.* (2017) Secreted metalloproteinase ADAMTS-3 inactivates reelin. *Journal of Neuroscience* **37**, 3181–91.

Purcell S., Neale B., Todd-Brown K. *et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Robinson J.T., Thorvaldsdottir H., Winckler W., Guttman M., Lander E.S., Getz G., Mesirov J.P. (2011) Integrative genomics viewer. *Nature Biotechnology* **29**, 24–6.

Schmutz S.M., Moker J.S., Clark E.G., Orr J.P. (1996) Chromosomal aneuploidy associated with spontaneous abortions and neonatal losses in cattle. *Journal of Veterinary Diagnostic Investigation* **8**, 91–5.

Sirmaci A., Erbek S., Price J. *et al.* (2010) A truncating mutation in SERPINB6 is associated with autosomal-recessive nonsyndromic sensorineural hearing loss. *The American Journal of Human Genetics* **86**, 797–804.

Švara T., Cociancich V., Šest K., Gombac M., Paller T., Starič J., Drögemüller C. (2016) Pulmonary hypoplasia and anasarca syndrome in Cika cattle. *Acta Veterinaria Scandinavica* **58**, 36.

Tang B.L., Hong W. (1999) ADAMTS: a novel family of proteases with an ADAM protease domain and thrombospondin 1 repeats. *FEBS Letters* **445**, 223–5.

Turner S.D. (2014) qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv*, R package version 0.1, 2. https://doi.org/10.1101/005165

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M., Rozen S.G. (2012) Primer3–new capabilities and interfaces. *Nucleic Acids Research* **40**, e115.

Whitlock B.K., Coffman E.A. (2015) Heritable congenital defects in cattle. In:*Bovine Reproduction* (Ed. by R.M. Hopper), pp.609–19. Wiley, Hoboken, NJ.

Whitlock B.K., Kaiser L., Maxwell H.S. (2008) Heritable bovine fetal abnormalities. *Theriogenology* **70**, 535–49.

Windsor P.A., Cavanagh J.A., Tammen I. (2006) Hydrops fetalis associated with pulmonary hypoplasia in Dexter calves. *Australian Veterinary Journal* **84**, 278–81.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Accession numbers of the publicly available genome sequences, which are used as control genomes

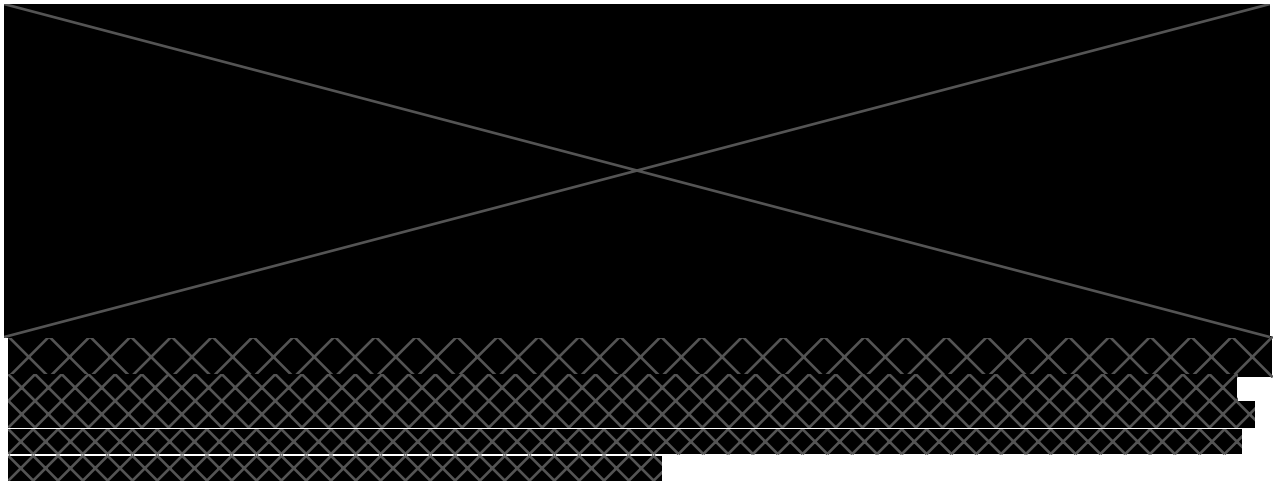
Table S2 Regions of ROH

Table S3 Genotyping of four candidate variants in Cika cattle

Table S4 *Post-hoc* Tukey test for pair-wise comparison of the average chromosomal read depth

Constitutional trisomy 20 in an aborted Holstein fetus with pulmonary hypoplasia and anasarca syndrome

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, visualizations, writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.13008
John Wiley and Sons License Number:	5236010339149



doi: 10.1111/age.13008

Constitutional trisomy 20 in an aborted Holstein fetus with pulmonary hypoplasia and anasarca syndrome

Irene Monika Häfliger*, Jørgen Steen Agerholm† and Cord Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland; †Department of Veterinary Clinical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Frederiksberg C 1870, Denmark

Accepted for publication 13 September 2020

Background

Pulmonary hypoplasia and anasarca (PHA) syndrome is a known congenital lethal disorder in cattle (OMIA 001562-9913) occurring sporadically in a variety of breeds. Typical signs of PHA include severe widespread subcutaneous edema and pulmonary hypoplasia or aplasia.[1,2] The extensive accumulation of fluid in the subcutaneous tissue and formation of large subcutaneous cysts deteriorate the body conformation, thus potentially causing dystocia. In cattle, PHA was recently described as either a recessive disorder owing to a missense variant in *ADAMTS3* in Cika cattle or as a dominant disorder caused by a trisomy of chromosome 20 in a single PHA-affected Holstein calf.[3]

Own analysis

A purebred Danish Holstein heifer started to abort a fetus on gestation day 252, but Cesarean section had to be performed as the fetus could not be delivered vaginally. The

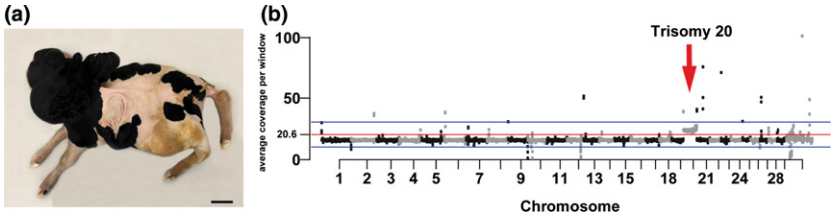


Figure 1 Trisomy 20 in a pulmonary hypoplasia and anasarca-affected Holstein fetus. (a) The fetus showed retarded development with a short-haired sparse haircoat not normal for a 252-day-old fetus. The body weight was 33.0 kg of which around 8 kg consisted of fluid that had oozed from the fetus during transport and an additional >4 kg in the thorax and abdomen. The fetus had severe deformation of the head and body, lesions that hindered a vaginal delivery. Bar: 10 cm. (b) Manhattan plot displaying the read depth of 500 kb windows across chromosomes. The red line indicates the average read depth of 20.6. Note the elevated coverage across chromosome 20 indicating a trisomy (red arrow)

pregnancy was the result of insemination with imported semen of a purebred Canadian Holstein, which as at September 2020 has around 1300 offspring in Denmark. The female fetus (stillborn) underwent postmortem examination and lesions consistent with PHA were found (Fig. 1a; Fig. S1).

Genomic DNA extracted from a sample from the ear was whole-genome sequenced as described before and deposited publicly (www.ebi.ac.uk/ena, sample accession no. SAMEA6528903).[3] Variant calling revealed no evidence for the presence of protein-changing variants in the *ADAMTS3* gene; neither the previously described missense variant causing PHA in Cika cattle was found nor any other. In addition, genome-wide filtering for variants that were present only in the genome of the case and absent in the genomes of 493 controls (Table S1) resulted in 12 heterozygous protein-changing variants of which four were absent in the variant catalog of run 7 of the 1000 Bull Genomes project (Table S2).[4] However, it was not evaluated further whether these four private variants represent *de novo* sequence variants or rare variants.

To check for chromosomal aberrations the read depth of 500 kb windows across chromosomes was calculated and plotted in a Manhattan plot (Fig. 1b). To compare the average read depth per chromosome QUALIMAP software was used.[5] Whereas the chromosomes 1–19, 21–29 and the X-chromosome had a read depth ranging from 18.3× to 20.1×, chromosome 20 had an average read depth of 28.2× (Fig. 1b). This 1.5-fold elevated read depth of chromosome 20 indicates a trisomy.

Comments

We suggest that the additional chromosome 20 in the Holstein fetus caused PHA. As there had not been reported any other PHA case in this Holstein family, the constitutional trisomy probably arose through non-disjunction during early cell division in the fertilized egg. Rare cases of chromosomal aberrations in cattle are described for a small number of chromosomes (18, 20, 21, 22, 28 and 29) and are associated with brachygnathia, blindness, reduced growth, aplasia of various organs and other malformations.[3,6–13] Whereas most of them are associated with severe body malformations and are lethal,[3,6–8,11,13] a few trisomies have been observed in living animals.[9,10,12] The absence of reports of most chromosomal aberrations indicates their lethality, most probably during early gestation.[11] This further case of chromosome 20 trisomy shared important features with a previously described association of trisomy 20 with PHA in cattle.[3] Therefore, it presents a recurrent mutation event apparently leading to bovine PHA, a disorder that was for a long time assumed to be recessively inherited.

Acknowledgements: We are grateful to Skovbjerg Dyrlægeteam for submitting the case and the farmer for donating it for research. We thank the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome re-sequencing experiments and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computing infrastructure. This study was partially funded by the Swiss National Science Foundation and the Danish Bovine Genetic Disease Programme.References1Windsor B.K. *et al.* (2006) Aust Vet J 84, 278–81.2Agerholm J.S. & Arnbjerg J. (2011) Vet Record 168, 190.3Häfliger I.M. *et al.* (2020) Anim Genet 51, 382–90.4Hayes B.D. & Daetwyler H.D. (2019) Annu Rev Anim Biosci 7, 89–102.5Okonechnikov K. *et al.* (2016) Bioinformatics 32, 292–4.6Agerholm J.S. & Christensen K. (1993) J Vet Med 40, 576–81.7Gallagher D.S. *et al.* (1999) Vet Pathol 36, 448–51.8Herzog A. *et al.* (1977) Ann Genet Sel Anim 9, 471–91.9Iannuzzi L. *et al.* (2001) Hereditas 134, 147–51.10Mayr B. *et al.* (1985) Cytogenet Cell Genet 39, 77–9.11Schmutz S.M. *et al.* (1996) J Vet Diagn Invest 8, 91–5.12Lioi M.B. *et al.* (1995) Genet Sel Evol 27, 473–6.13Häfliger I.M. *et al.* (2020) Anim Genet 51, 483–4.

Correspondence: Chuzhao Lei (leichuzhao1118@126.com) and Bizhi Huang (hbz@ynbp.cn)

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Necropsy findings of the PHA-affected Holstein fetus

Table S1. Accession nos of the 493 publicly available genome sequences of various breeds that were sequenced in the course of other ongoing studies used as control genomes

Table S2. Private protein-changing heterozygous sequence variants in the affected fetus and their distribution in the 1000 Bull Genomes data

A heterozygous missense variant in *MAP2K2* in a stillborn Romagnola calf with skeletal-cardio-enteric dysplasia

Journal:	Animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation, support genetic analyzes, writing – review and editing
Displayed version:	published version
DOI:	10.3390/ani11071931

animals

MDPI

Article

A Heterozygous Missense Variant in *MAP2K2* in a Stillborn Romagnola Calf with Skeletal-Cardio-Enteric Dysplasia

Joana G. P. Jacinto [1,2], Irene M. Häfliger [2], Arcangelo Gentile [1] and Cord Drögemüller [2,*]

[1] Department of Veterinary Medical Sciences, University of Bologna, Ozzano Emilia, 40064 Bologna, Italy; joana.goncalves2@studio.unibo.it (J.G.P.J.); arcangelo.gentile@unibo.it (A.G.)
[2] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch

Simple Summary: Skeletal dysplasias encompass a clinical-, pathological- and genetically heterogeneous group of disorders characterized by abnormal cartilage and/or bone formation, growth, and remodeling. They may belong to the so-called RASopathies, congenital conditions caused by heterozygous variants in genes that encode components of the Ras/mitogen-activated protein kinase (MAPK) cell signaling pathway. Herein, an affected calf of the Italian Romagnola breed was reported showing a skeletal-cardio-enteric dysplasia. We identified a most likely disease-causing mutation in the *MAP2K2* gene by whole-genome sequencing (WGS). The *MAP2K2* gene is known to be related with dominant inherited cardio-facio-cutaneous syndrome in humans, but it was so far unknown to cause a similar disease in domestic animals. We assume that the identified missense variant that was predicted to impair the function of the protein, occurred either within the germline of the dam or post-zygotically in the embryo. Rare lethal diseases such as the skeletal-cardio-enteric dysplasia in livestock are usually not characterized to the molecular level, mainly because of the lack of funds and diagnostic opportunities. Precise WGS-based diagnostics enables the understanding of rare diseases and supports the value of monitoring cattle breeding populations for fatal genetic defects.

Abstract: RASopathies are a group of developmental disorders caused by dominant mutations in genes that encode components of the Ras/mitogen-activated protein kinase (MAPK) cell signaling pathway. The goal of this study was to characterize the pathological phenotype of a Romagnola stillborn calf with skeletal-cardio-enteric dysplasia and to identify a genetic cause by whole-genome sequencing (WGS). The calf showed reduced fetal growth, a short-spine, a long and narrow face, cardiac defects and heterotopy of the spiral colon. Genetic analysis revealed a private heterozygous missense variant in *MAP2K2*:p.Arg179Trp, located in the protein kinase domain in the calf, and not found in more than 4500 control genomes including its sire. The identified variant affecting a conserved residue was predicted to be deleterious and most likely occurred de novo. This represents the first example of a dominant acting, and most likely pathogenic, variant in *MAP2K2* in domestic animals, thereby providing the first *MAP2K2*-related large animal model, especially in respect to the enteric malformation. In addition, this study demonstrates the utility of WGS-based precise diagnostics for understanding sporadic congenital syndromic anomalies in cattle and the general utility of continuous surveillance for rare hereditary defects in cattle.

Keywords: cattle; cardiac defect; development; congenital malformations; heterotopy; precision medicine; short spine; RASopathy

Citation: Jacinto, J.G.P.; Häfliger, I.M.; Gentile, A.; Drögemüller, C. A Heterozygous Missense Variant in *MAP2K2* in a Stillborn Romagnola Calf with Skeletal-Cardio-Enteric Dysplasia. *Animals* **2021**, *11*, 1931. https://doi.org/10.3390/ani11071931

Academic Editors: Francesca Ciotola, Sara Albarella and Vincenzo Peretti

Received: 5 May 2021
Accepted: 25 June 2021
Published: 29 June 2021

1. Introduction

Genetic skeletal dysplasias encompass a clinical-, pathological-, and genetically heterogeneous group of rare disorders characterized by abnormal cartilage and/or bone formation, growth, and remodeling [1]. In human medicine, 461 different skeletal disorders

are classified into 42 subtypes [2]. At the present in humans, pathogenic variants affecting more than 437 different genes have been found to be associated with these disorders [2]. In veterinary medicine, genetic skeletal dysplasias are not classified in such detail. Nevertheless, with the progressively widespread availability of molecular tools for genetic mapping, such as single-nucleotide polymorphism (SNP) arrays, and for mutation analysis, such as short-read based whole-genome sequencing (WGS), the recognition of disease-causing pathogenic variants has drastically improved [3,4]. In fact, the 1000 Bull Genomes Project now encompasses a genetic variation from over 4100 genomes providing a comprehensive database for the imputation of genetic polymorphisms for genomic prediction in all cattle breeds, improving the accuracy of genomic prediction in the identification of causal mutations [5]. The OMIA (Online Mendelian Inheritance of Animals) currently lists 22 skeletal disorders in cattle with a known causal mutation, e.g., recessively inherited mostly breed-specific disorders such as the brachyspina syndrome in Holstein (OMIA 000151-9913) [6] and the paunch calf syndrome in Romagnola [7] and Marchigiana [8] (OMIA 001722-9913), or dominantly inherited disorders such as the bovine achondrogenesis type II (OMIA 001926-9913) [9], or cases of facial dysplasia in the progeny of a single bull (OMIA 002090-9913) [10]. The latter two diseases have been shown to result from de novo mutation events in the paternal germline.

This study aimed to characterize in detail the pathological phenotype of a Romagnola calf with skeletal cardio-enteric dysplasia, and to discover a genetic variant causing the abnormality by WGS.

2. Materials and Methods

2.1. Pathological Investigation

A stillborn Romagnola male calf was referred to the Department of Veterinary Medical Sciences, University of Bologna for post-mortem examination. The calf resulted from insemination with semen from a purebred Romagnola bull on a Romagnola cow. The pedigree of both parents showed no common ancestor within four generations. The truncal length was measured from the occipital bone to the tuber coxae. Additionally, radiographic images of the spine were obtained before starting necropsy.

2.2. DNA Samples

Genomic DNA was isolated from skin and cartilage from the ear of the calf and from the semen of the sire using Promega Maxwell RSC DNA system (Promega, Dübendorf, Switzerland). In addition, genomic DNA samples from EDTA-blood samples of 100 apparently normal Romagnola cattle were extracted with the same methodology and used as controls.

2.3. Whole-Genome Sequencing

The genome of the affected calf was sequenced as described before resulting in an average read coverage of approximately 17.4× [11]. Single-nucleotide variants (SNVs) and small indel variants were called subsequently as reported earlier [5], except for the trimming, which was carried out using fastp [12]. Further data processing was carried out according to Häfliger et al., 2020 [13]. The impact of the called sequence variants was evaluated with snpeff v4.3 [14], using NCBI Annotation Release 106 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/; accessed on 30 April 2021). In order to search for private variants, we compared the genotypes of the affected calf with 598 bovine genomes of different breeds that have been sequenced in other ongoing studies and are freely available (Table S1) in the European Nucleotide Archive (SAMEA7690195 is the sample accession number of the affected calf; http://www.ebi.ac.uk/en; acceded on 30 April 2021). The occurrence of these variants was then investigated in a global control cohort of 4110 genomes of different breeds (1000 Bull Genomes Project run 8; www.1000bullgenomes.com; acceded on 30 April 2021) [5]. Integrative Genomics Viewer

(IGV) [15] software version 2.0 was used to manually look at genomic regions where possible candidate genes map.

2.4. Targeted Genotyping

PCRs were carried out using Amplitaq Gold Master Mix (Thermofisher, Rotkreuz, Switzerland). The subsequent bi-directional sequencing of PCR products was carried out after shrimp alkaline phosphatase (Roche, Basel, Switzerland) and exonuclease I (NEB, Axon lab, Baden, Switzerland) incubation using the PCR primers with the ABI BigDye Terminator Sequencing Kit 3.1 (Applied Biosystems, Zug, Switzerland) on an ABI 3730 capillary sequencer (Applied Biosystems). The *MAP2K2* missense variant (NM_001038071.2: g.19923991C>T) was genotyped using the following primers: 5′-GGCTTAACAGAGGATGCCCC-3′ (forward primer) and 5′-CTGGAAAACCTGGAAA-TCGGG-3′ (reverse primer). The evaluation of the sequence data was carried out with Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA).

2.5. Evaluation of the Molecular Consequences of Amino Acid Substitutions

PROVEAN [16] and PredictSNP1 [17] were used to predict the biological consequences of the discovered variant on protein. For multispecies sequence alignments the following NCBI proteins accessions were used: NP_001033160.2 (*Bos taurus*), NP_109587.1 (*Homo sapiens*), XP_003318859.1 (*Pan troglodytes*), XP_001118016.2 *(Macaca mulatta*), NP_001041601.1 (*Canis lupus*), NP_075627.2 (*Mus musculus*), NP_579817.1 (*Rattus norvegicus*), NP_990719.1 (*Gallus gallus*), NP_001032468.2 (*Danio rerio*).

2.6. Sequence Accessions

Genomic positions in the cow genome refers to the the ARS-UCD1.2 assembly. All references to the bovine *MAP2K2* gene correspond to the NCBI accessions NC_037334.1 (chromosome 7, ARS-UCD1.2), NM_001038071.2 (*MAP2K2* mRNA), and NP_001033160.2 (*MAP2K2* protein). For the *MAP2K2* protein, the Uniprot database (https://www.uniprot.org/) (accessed on 15 March 2021) accession number A0A3S5ZPX3 was used.

3. Results

3.1. Pathological Phenotype

At birth, the calf weighed only 20.3 kg (normal weight of Romagnola calves at birth 40 Kg–mean) and its trunk was disproportionately short for the body size and legs (Figure 1; Figure S1). Shortening of the neck was very pronounced, and it looked as though the head was fixed to the chest. The length of trunk, as measured from the occipital bone to the point of the buttock, was approximately 40 cm. Moreover, it displayed a mild kyphosis at the level of the thoracolumbar region. The limbs were 70 cm long and slender (dolichostenomelia).

Facial deformities were characterized by narrow, longer and laterally deviated splanchnocranium. The lower jaw was slightly longer than normal.

Radiological examination of the axial skeleton showed only a reduced size of the vertebral bodies, but no shape abnormalities (Figure S2).

At gross pathology, the examination of the abdominal cavity showed the absence of the omentum and heterotopy of the spiral colon, the latter characterized by a complete detachment and complete displacement of the spiral loop of the ascending colon from the mesojejunum (Figure 2a,b).

The calf also showed an evident 2 cm diameter persistent patent ductus and pathological cardiac abnormalities, including globous shape, enlarged right ventricle and pulmonic stenosis.

Based on these pathological observations, the animal was considered to present a skeletal-cardio-enteric dysplasia.

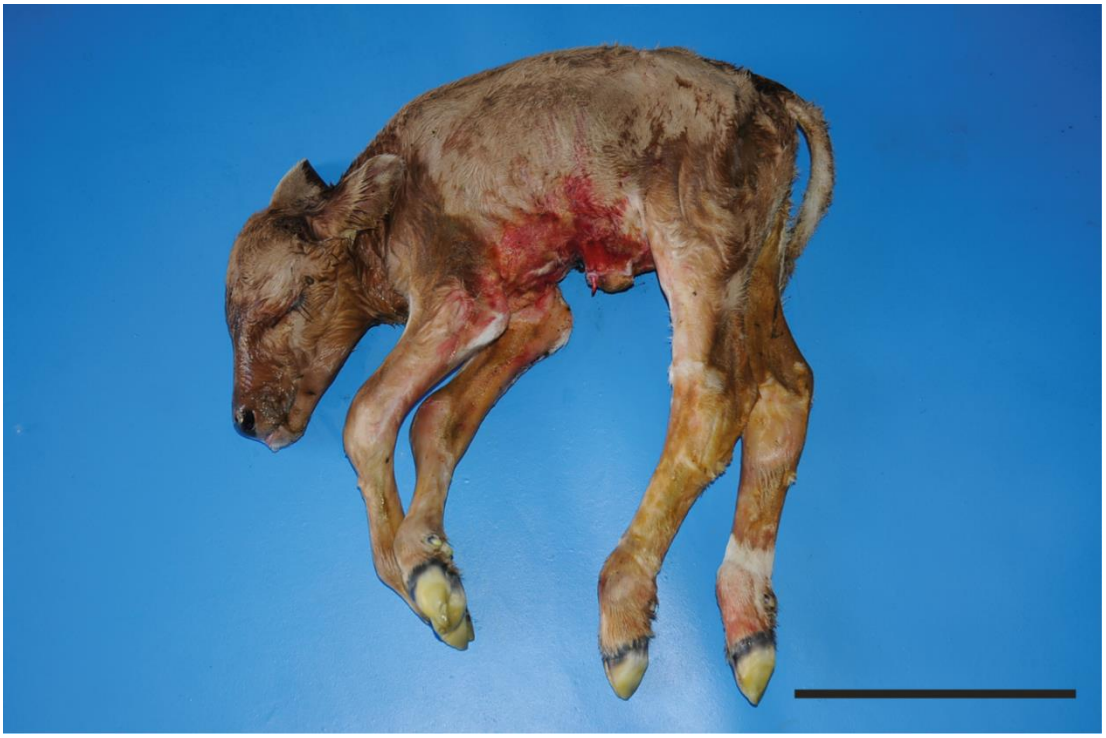


Figure 1. Stillborn Romagnola calf with skeletal-cardio-enteric dysplasia: Strikingly, the length of the spine appeared disproportionately short for the height and legs. Mild kyphosis of the thoraco-lumbar vertebral column is also evident. Bar, 30 cm.

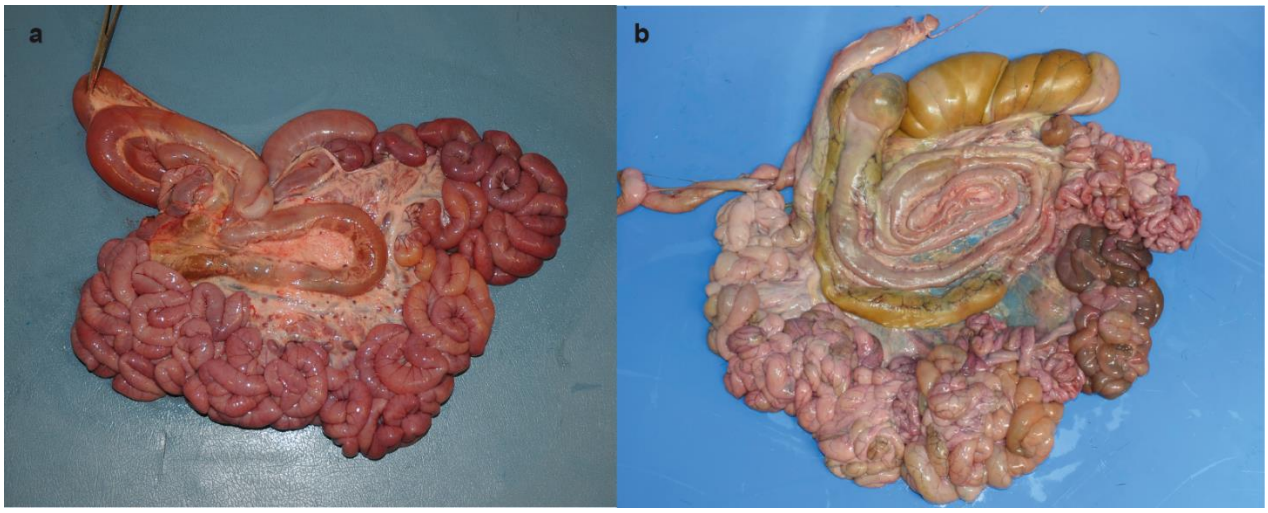


Figure 2. Topography of the spiral loop of the ascending colon. (**a**) Topography of the spiral loop of the ascending colon in the stillborn Romagnola calf with skeletal-cardio-enteric dysplasia. Note the complete displacement of the spiral loop of the ascending colon from the mesojejunum. (**b**) Topography of the spiral loop of the ascending colon in a control. Note the spiral loop of the ascending colon located at the middle of the mesojejunum.

3.2. Genetic Analysis

Assuming a spontaneous mutation as the most likely cause for this congenital malformation, the whole genome of the affected calf was sequenced. To evaluate the presence of coding protein-changing variants, filtering of WGS for private variants present in the calf and not present in the 598 control genomes was performed. This approach identified 381 heterozygous private protein-changing variants predicted to have moderate or severe alteration on the encoded proteins. These variants were then tested for their presence

in a global control cohort of 4110 genomes of various breeds collected in run 8 of the 1000 Bull Genomes Project [5], which revealed 66 protein-changing variants only occurring in heterozygous state in the genome of the affected calf. These 66 variants were subsequently evaluated using IGV software, confirming 63 as true variants (Table S2). Of all the remaining private variants, one occurred in a possible candidate gene for observed phenotype (Figure 3a). The heterozygous variant at chr7:19923991C>T represents a missense variant in exon 5 of the *MAP2K2* gene (NM_001038071.2: c.535C>T; Figure 3c). This variant alters the amino acid of the *MAP2K2* protein at site 179 (NP_001033160.2: p.Arg179Trp) located in the protein kinase domain (Figure 3e). Furthermore, the arginine to tryptophan substitution affects an evolutionary conserved residue (Figure 3f) and has been predicted to be harmful (PROVEAN score −4.708; Predict SNP score 61%). In order to confirm and finally evaluate the *MAP2K2* variant, the corresponding genome region was amplified by PCR and then analyzed by Sanger sequencing in the calf, its sire, and 100 controls of the same breed. Unfortunately, no biological sample of the dam that was slaughtered in-between was available. When analyzing the sequencing data, we found that the calf was indeed heterozygous for the *MAP2K2* variant detected, while the sire and another 100 controls from the Italian Romagnola population were homozygous for the wild-type allele (Figure 3b,c).

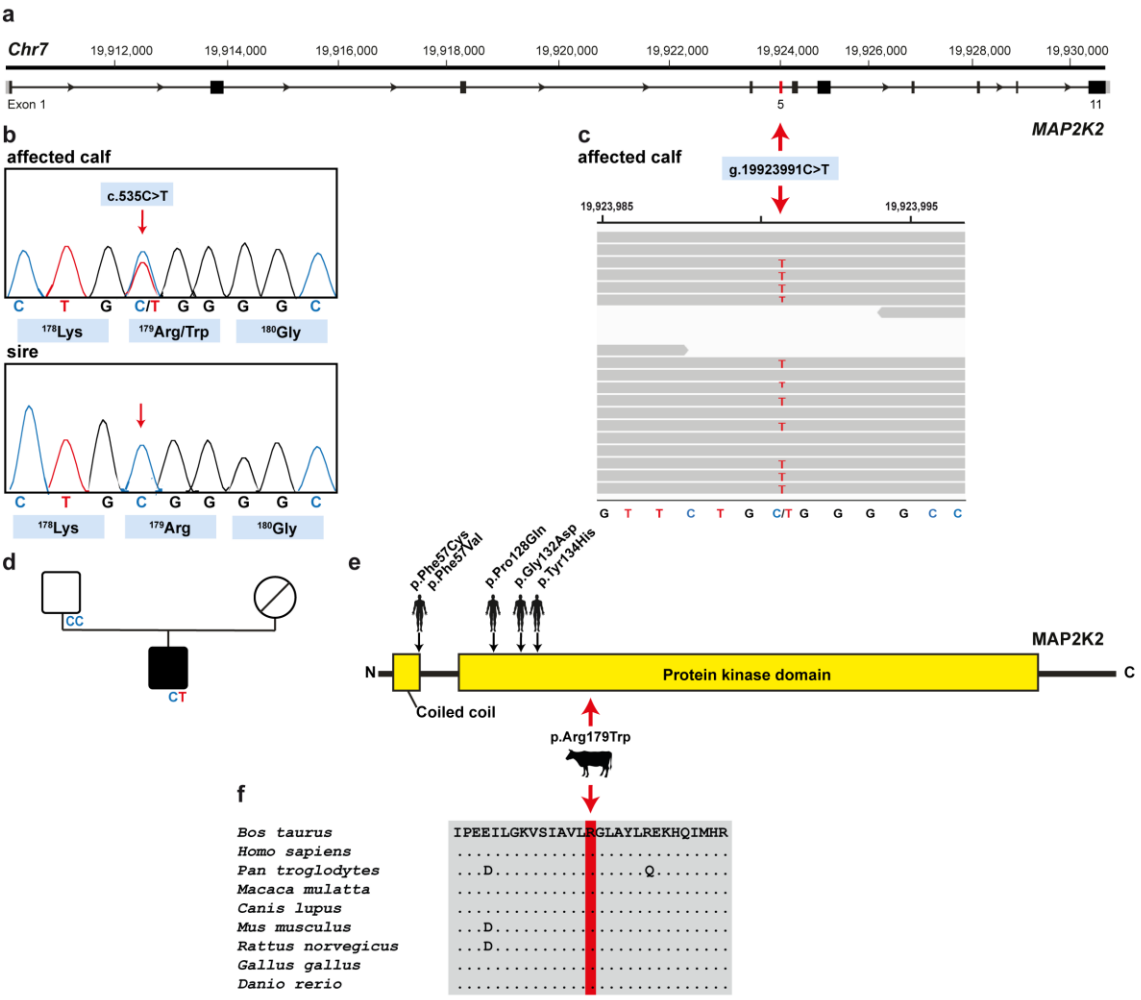


Figure 3. *MAP2K2* missense variant in a Romagnola calf with skeletal-cardio-enteric dysplasia: (**a**) *MAP2K2* gene structure representation of the exact position of the exon 5 variant on the chromosome 7 (red arrow); (**b**) Electropherograms confirming heterozygosity in the affected calf and the absence of the variant in the germline of the sire. (**c**) IGV screenshot presenting

the Chr7: g.19923991C>T variant in the calf. (**d**) Pedigree of the case. The affected male calf is represented with full black symbol, while both non-affected parents are represented by full white symbols. Unknown genotype is represented by a symbol with a diagonal line. (**e**) Schematic representation of *MAP2K2* protein and its functional domains and summary of known human *MAP2K2* mutations. The position of the mutation detected in the affected Romagnola calf is indicated by a red arrow, while know human *MAP2K2* mutations associated with cardio-facio-cutaneous syndrome are indicated by black arrows (OMIM601263). (**f**) Multiple sequence alignment of the protein kinase domain of the *MAP2K2* protein around the position of the p.Arg179Trp variant shows complete evolutionary conservation across all species.

In addition, when looking for homozygous variants, filtering revealed two private protein-changing variants present in the genome of the affected calf (Table S2).

4. Discussion

Non-infectious syndromic congenital malformations of newborns in cattle occur rarely and are most often not further diagnosed in detail. We have carried out a comprehensive pathological and genetic examination in a Romagnola stillborn calf, revealing a skeletal-cardio-enteric dysplasia. We then investigated the hypothesis of a spontaneous mutation as a possible reason for this congenital phenotype. Using state-of-the-art genetic approaches involving WGS, geneticist have only about a 50:50 chance of quickly identifying variants that are causal for developmental anomalies in humans [18]. So far, similar data is missing for veterinary medicine, mostly due to the lack of resources, although the scientific value for biomedical research is widely accepted [4]. Analysis of the genome sequence of the studied case identified a missense variant in a plausible candidate gene affecting the protein kinase domain of *MAP2K2*. In addition, this variant was only present in the genome of the affected calf, and did not occur in a global control cohort of more than 4500 bovine genomes of different breeds. Therefore, considering the rarity of this coding variant and the function of the *MAP2K2*, it was considered to represent the most likely genetic cause for the observed phenotype. To the best of our knowledge, no pathogenic variant in the *MAP2K2* gene has been reported in domestic species. Therefore, this study represents the first large animal model for a *MAP2K2*-related congenital skeletal disorder in cattle.

Furthermore, the PCR to detect the mutant allele in the sire using DNA extracted from semen showed a homozygous wild-type genotype. Therefore, we could exclude the father as a mosaic ancestor. However, to definitively prove that the identified variant in *MAP2K2* indeed occurred de novo, genotyping of the dam would be needed. Therefore, we speculate that the mutation either arose post-zygotically during fetal development of the affected calf or represents a maternally derived germline mutation.

The RAS/mitogen activated protein kinase (MAPK) cell signaling pathway plays an important role in the regulation of the cell cycle and differentiation [19]. In particular, during embryonic development, it represents one of the main pathways for the transduction of intracellular signals in response to all types of mitogens (e.g., growth factors), which initiates proliferation, survival, and anti-apoptotic programs [20]. Furthermore, the dysregulation of RAS/MAPK-dependent developmental processes has significant pathophysiological consequences [20].

In humans, somatic mutations leading to hyperactivation of the RAS/MAPK signaling cascade may cause cancers [21], whereas germline or de novo mutations in the developing embryo are responsible for several rare genetic conditions, collectively termed RASopathies. These disorders have common phenotypes, such as a short stature, heart defects, facial abnormalities, and cognitive impairments, often associated with abnormal central nervous system development, and include conditions such as neurofibromatosis type 1 (OMIM 162200) [22], Noonan (OMIM 163950) [23], LEOPARD (OMIM 151100) [24], Costello (OMIM 218040) [25], and cardio-facio-cutaneous (CFC; OMIM 115150) [26] syndromes. In humans, RASopathies are a highly heterogenetic group of genetic disorders, being associated to more than 20 causal genes [20]. These genes encode proteins that belong to, or regulate, the RAS/MAPK cell signaling pathway. Their mutations explain the pathophisological mechanisms, such as the abnormal development of various tissues (e.g., cardiac or craniofacial

defects), as well as the altered hormonal response and consequent endocrine dysfunctions (e.g., growth hormone insensitivity, and growth retardation) [20]. In particular, individuals affected by CFC show characteristic craniofacial dysmorphic features, short stature, cardiac defects, ectodermal anomalies, and developmental delay [26]. Currently, dominantly inherited mutations in four genes have been associated with CFC syndrome: *BRAF* [27], *MAP2K1* [28], *MAP2K2* [26,28], and *KRAS* [27]. Interestingly, the affected calf presented in this study revealed a dominant mutation in *MAP2K2*, and showed retarded growth, skeletal dimorphisms, including a long and narrow face and shortening of the vertebral column, and cardiac defects such as persistent patent ductus and pulmonary stenosis. These findings resemble the phenotype of human CFC syndrome. However, the animal did not show any alteration on the integumentary system.

Moreover, the calf revealed heterotopy of the spiral colon. In cattle, this anatomic anomaly may originate from the loose attachment of the spiral loops to the mesojejunum, abnormal elongation of the mesocolon, overlapping of adjacent loops, abnormal coiling, and finally, complete separation of the spiral loop of the ascending colon from the mesojejunum [29]. This condition can predispose to intestinal injury, such as volvulus and intussusception [29]. Similarly, in human medicine, the so-called intestinal malrotation is a congenital anomaly resulting from incomplete rotation and fixation of the intestine during embryological and fetal development, which may predispose to midgut volvulus and lead to duodenal obstruction and strangulation of the circulation in the superior mesenteric vessels [30]. In humans and mice, *MAP2K2* is known to be highly expressed in the colon and duodenum [31,32]. However, mutations in *MAP2K2* have so far not been associated with enteric congenital defects.

Furthermore, the deleterious nature of the variant and the conservation of the affected arginine amino acid residue of *MAP2K2* at position 179 in protein kinase domain also suggest that this variant is most likely pathogenic. Indeed, mutations of *MAP2K2* affect the negative regulatory region or the core catalytic domain of the kinases, leading to increased kinase activity and gain-of-function effect on the RAS/MAPK pathway [28,33,34].

5. Conclusions

The investigation of this case enabled a pathological and molecular genetic study that, for the first time, allowed the diagnosis of a dominantly inherited skeletal-cardio-enteric dysplasia in a calf associated with a *MAP2K2* missense variant. Thus, we hereby present the first large animal model for similar human diseases. Furthermore, this example highlights the usefulness of genetically precise diagnosis for understanding sporadic cases of congenital disorders caused by de novo mutations, and the need for continuous monitoring of genetic lethal disorders in cattle breeding. Additionally, novel discoveries in large animals such as cattle are useful as biomedical models for human diseases.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ani11071931/s1, Table S1: EBI Accession numbers of all publicly available genome sequences. We compared the genotypes of the calf with 598 cattle genomes of various breeds that had been sequenced in the course of other ongoing studies and that were publicly available, Table S2: List of the remaining variants after the comparison to the global control cohort of 4110 genomes of other breeds (1000 Bull Genomes Project run 8; www.1000bullgenomes.com; acceded on 30 April 2021) and after IGV visual inspections, revealing 63 protein-changing variants with a predicted moderate or high impact only present in the affected calf. Figure S1: Newborn Romagnola healthy calf. Figure S2: Radiographic image of the thoraco-lumbar region of the affected calf.

Author Contributions: Conceptualization, C.D. and A.G.; methodology, I.M.H., A.G. and C.D.; validation, J.G.P.J.; formal analysis, J.G.P.J., I.M.H., A.G. and C.D.; investigation, J.G.P.J. and A.G.; resources, C.D.; data curation, I.M.H.; writing—original draft preparation, J.G.P.J.; writing—review and editing, J.G.P.J., I.M.H., A.G. and C.D.; visualization, A.G. and J.G.P.J.; supervision, C.D. and A.G.; project administration, C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially funded by the Swiss National Science Foundation.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but part of a clinical and pathological veterinary diagnostic case.

Informed Consent Statement: Not applicable.

Data Availability Statement: The whole-genome data of the affected calf (sample ID PCS1780) is freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA7690195.

Acknowledgments: We are grateful to Nathalie Besuchet Schmutz for expert technical assistance and the Interfaculty Bioinformatics Unit of the University of Bern for providing high-performance computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Song, D.; Maher, C.O. Spinal Disorders Associated with Skeletal Dysplasias and Syndromes. *Neurosurg. Clin. N. Am.* **2007**, *18*, 499–514. [CrossRef] [PubMed]
2. Mortier, G.R.; Cohn, D.H.; Cormier-Daire, V.; Hall, C.; Krakow, D.; Mundlos, S.; Nishimura, G.; Robertson, S.; Sangiorgi, L.; Savarirayan, R.; et al. Nosology and classification of genetic skeletal disorders: 2019 revision. *Am. J. Med. Genet. Part A* **2019**, *179*, 2393–2419. [CrossRef]
3. Dittmer, K.E.; Thompson, K.G. Approach to Investigating Congenital Skeletal Abnormalities in Livestock. *Vet. Pathol.* **2015**, *52*, 851–861. [CrossRef] [PubMed]
4. Lyons, L.A. Precision medicine in cats—The right biomedical model may not be the mouse! *PLoS Genet.* **2020**, *16*, 8–11. [CrossRef] [PubMed]
5. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef]
6. Charlier, C.; Agerholm, J.S.; Coppieters, W.; Karlskov-Mortensen, P.; Li, W.; de Jong, G.; Fasquelle, C.; Karim, L.; Cirera, S.; Cambisano, N.; et al. A Deletion in the Bovine FANCI Gene Compromises Fertility by Causing Fetal Death and Brachyspina. *PLoS ONE* **2012**, *7*, 2–8. [CrossRef]
7. Testoni, S.; Bartolone, E.; Rossi, M.; Patrignani, A.; Bruggmann, R.; Lichtner, P.; Tetens, J.; Gentile, A.; Drögemüller, C. KDM2B Is Implicated in Bovine Lethal Multi-Organic Developmental Dysplasia. *PLoS ONE* **2012**, *7*. [CrossRef]
8. Murgiano, L.; Militerno, G.; Sbarra, F.; Drögemüller, C.; Jacinto, J.G.P.; Gentile, A.; Bolcato, M. KDM2B-associated paunch calf syndrome in Marchigiana cattle. *J. Vet. Intern. Med.* **2020**, *34*, 1657–1661. [CrossRef]
9. Jacinto, J.G.P.; Häfliger, I.M.; Letko, A.; Drögemüller, C.; Agerholm, J.S. A large deletion in the COL2A1 gene expands the spectrum of pathogenic variants causing bulldog calf syndrome in cattle. *Acta Vet. Scand.* **2020**, *62*, 49. [CrossRef]
10. Agerholm, J.S.; McEvoy, F.J.; Heegaard, S.; Charlier, C.; Jagannathan, V.; Drögemüller, C. A de novo missense mutation of FGFR2 causes facial dysplasia syndrome in Holstein cattle. *BMC Genet.* **2017**, *18*. [CrossRef]
11. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, 1–9. [CrossRef]
12. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [CrossRef] [PubMed]
13. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef] [PubMed]
14. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]
15. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant review with the integrative genomics viewer. *Cancer Res.* **2017**, *77*, e31–e34. [CrossRef]
16. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]
17. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Comput. Biol.* **2014**, *10*. [CrossRef] [PubMed]
18. Kingsmore, S.F.; Cakici, J.A.; Clark, M.M.; Gaughran, M.; Feddock, M.; Batalov, S.; Bainbridge, M.N.; Carroll, J.; Caylor, S.A.; Clarke, C.; et al. A Randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. *Am. J. Hum. Genet.* **2019**, *105*, 719–733. [CrossRef] [PubMed]

19. Takenouchi, T.; Shimizu, A.; Torii, C.; Kosaki, R.; Takahashi, T.; Saya, H.; Kosaki, K. Multiple café au lait spots in familial patients with *MAP2K2* mutation. *Am. J. Med. Genet. Part A* **2014**, *164*, 392–396. [CrossRef]
20. Tajan, M.; Paccoud, R.; Branka, S.; Edouard, T.; Yart, A. The RASopathy family: Consequences of germline activation of the RAS/MAPK pathway. *Endocr. Rev.* **2018**, *39*, 676–700. [CrossRef]
21. Santarpia, L.; Lippman, S.M.; El-Naggar, A.K. Targeting the MAPK–RAS–RAF signaling pathway in cancer therapy. *Expert Opin. Ther. Targets* **2012**, *16*, 103–119. [CrossRef] [PubMed]
22. Upadhyaya, M.; Majounie, E.; Thompson, P.; Han, S.; Consoli, C.; Krawczak, M.; Cordeiro, I.; Cooper, D.N. Three different pathological lesions in the NF1 gene originating de novo in a family with neurofibromatosis type 1. *Hum. Genet.* **2003**, *112*, 12–17. [CrossRef] [PubMed]
23. Tartaglia, M.; Kalidas, K.; Shaw, A.; Song, X.; Musat, D.L.; Van der Burgt, I.; Brunner, H.G.; Bertola, D.R.; Crosby, A.; Ion, A.; et al. PTPN11 mutations in noonan syndrome: Molecular spectrum, genotype-phenotype correlation, and phenotypic heterogeneity. *Am. J. Hum. Genet.* **2002**, *70*, 1555–1563. [CrossRef] [PubMed]
24. Kalidas, K.; Shaw, A.C.; Crosby, A.H.; Newbury-Ecob, R.; Greenhalgh, L.; Temple, I.K.; Law, C.; Patel, A.; Patton, M.A.; Jeffery, S. Genetic heterogeneity in LEOPARD syndrome: Two families with no mutations in PTPN11. *J. Hum. Genet.* **2005**, *50*, 21–25. [CrossRef] [PubMed]
25. Kerr, B.; Allanson, J.; Delrue, M.A.; Gripp, K.W.; Lacombe, D.; Lin, A.E.; Rauen, K.A. The diagnosis of Costello syndrome: Nomenclature in Ras/MAPK pathway disorders. *Am. J. Med. Genet. A* **2008**, *146A*, 1218–1220. [CrossRef]
26. Rauen, K.A.; Tidyman, W.E.; Estep, A.L.; Sampath, S.; Peltier, H.M.; Bale, S.J.; Lacassie, Y. Molecular and functional analysis of a novel MEK2 mutation in cardio-facio-cutaneous syndrome: Transmission through four generations. *Am. J. Med. Genet. A* **2010**, *152A*, 807–814. [CrossRef]
27. Niihori, T.; Aoki, Y.; Narumi, Y.; Neri, G.; Cavé, H.; Verloes, A.; Okamoto, N.; Hennekam, R.C.M.; Gillessen-Kaesbach, G.; Wieczorek, D.; et al. Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome. *Nat. Genet.* **2006**, *38*, 294–296. [CrossRef]
28. Rodriguez-Viciana, P.; Tetsu, O.; Tidyman, W.E.; Estep, A.L.; Conger, B.A.; Cruz, M.S.; McCormick, F.; Rauen, K.A. Germline mutations in genes within the MAPK pathway cause cardio-facio-cutaneous syndrome. *Science* **2006**, *311*, 1287–1290. [CrossRef]
29. Gentile, A.; Bolcato, M.; Militerno, G.; Rademacher, G.; Desrochers, A.; Grandis, A. Heterotopy ("Error loci") of the spiral loop of the ascending colon in cattle. *PLoS ONE* **2019**, *14*. [CrossRef]
30. Salehi Karlslätt, K.; Pettersson, M.; Jäntti, N.; Szafranski, P.; Wester, T.; Husberg, B.; Ullberg, U.; Stankiewicz, P.; Nordgren, A.; Lundin, J.; et al. Rare copy number variants contribute pathogenic alleles in patients with intestinal malrotation. *Mol. Genet. Genomic Med.* **2019**, *7*, 1–9. [CrossRef]
31. Fagerberg, L.; Hallström, B.M.; Oksvold, P.; Kampf, C.; Djureinovic, D.; Odeberg, J.; Habuka, M.; Tahmasebpoor, S.; Danielsson, A.; Edlund, K.; et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **2014**, *13*, 397–406. [CrossRef]
32. Yue, F.; Cheng, Y.; Breschi, A.; Vierstra, J.; Wu, W.; Ryba, T.; Sandstrom, R.; Ma, Z.; Davis, C.; Pope, B.D.; et al. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **2014**, *515*, 355–364. [CrossRef]
33. Dentici, M.L.; Sarkozy, A.; Pantaleoni, F.; Carta, C.; Lepri, F.; Ferese, R.; Cordeddu, V.; Martinelli, S.; Briuglia, S.; Digilio, M.C.; et al. Spectrum of MEK1 and MEK2 gene mutations in cardio-facio-cutaneous syndrome and genotype-phenotype correlations. *Eur. J. Hum. Genet.* **2009**, *17*, 733–740. [CrossRef] [PubMed]
34. Anastasaki, C.; Estep, A.L.; Marais, R.; Rauen, K.A.; Patton, E.E. Kinase-activating and kinase-impaired cardio-facio-cutaneous syndrome alleles have activity during zebrafish development and are sensitive to small molecule inhibitors. *Hum. Mol. Genet.* **2009**, *18*, 2543–2554. [CrossRef] [PubMed]

APOB-associated cholesterol deficiency in Holstein cattle is not a simple recessive disease

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	statistical analyzes, visualizations, writing original draft and revisions
Displayed version:	published version
DOI:	10.1111/age.12801
John Wiley and Sons License Number:	5236041001712

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12801

APOB-associated cholesterol deficiency in Holstein cattle is not a simple recessive disease

Irene Monika Häfliger*, Sonja Hofstetter*, Thomas Mock†, Manuela Hanna Stettler†, Mireille Meylan†, Kemal Mehinagic‡, Nadine Stokar-Regenscheit‡ and Cord Drögemüller*

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland . †Clinic for Ruminants, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland . ‡Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland .

Summary

In 2015, cholesterol deficiency (CD) was reported for the first time as a new recessive defect in Holstein cattle. After GWAS mapping and identification of a disease-associated haplotype, a causative loss-of-function variant in *APOB* was identified. CD-clinically affected *APOB* homozygotes showed poor development, intermittent diarrhea and hypocholesterolemia and, consequently, a limited life expectation. Herein, we present a collection of 18 cases clinically diagnosed as CD-affected *APOB* heterozygotes. CD-clinically affected heterozygotes show reduced cholesterol and triglyceride blood concentrations. The differences in total blood cholesterol and triglycerides between nine CD-clinically affected and 36 non-affected heterozygotes were significant. As only some *APOB* heterozygotes show the clinical CD phenotype, we assume that the penetrance is reduced in heterozygotes compared to the fully penetrant effect observed in homozygotes. We conclude that *APOB*-associated CD represents most likely an incomplete dominant inherited metabolic disease with incomplete penetrance in heterozygotes.

Keywords apolipoprotein B, calf survival, genetic disorder, incomplete penetrance, rare disease

Recently, numerous calves showing retarded growth and intermittent diarrhea, accompanied by pronounced hypocholesterolemia and low triglyceride concentrations, were observed in the global Holstein population (Kipp *et al.* 2016; Mock *et al.* 2016; Schütz *et al.* 2016). This previously unknown fat metabolism disorder was termed cholesterol deficiency (CD; OMIA 001965-9913). After a genome-wide association study uncovering a disease-associated haplotype on chromosome 11 (Kipp *et al.* 2016), a 1.3-kb insertion of a transposable LTR element in the coding sequence of the bovine *APOB* gene was identified as a disease-causing variant (Menzi *et al.* 2016). These findings were independently confirmed (Charlier 2016; Schütz *et al.* 2016). So far, the defect is clinically identifiable only in the homozygote form: Kipp *et al.* (2016) reported nine cases homozygous for the disease-associated haplotype, and Duff *et al.* (2016), Mock *et al.* (2016) and Schütz *et al.* (2016) reported one, six and nine *APOB* homozygote mutant cases respectively. The disease-associated haplotype carrying the *APOB* variant traces back to the North American bull, Maughlin Storm, born in 1991 and used extensively in the Holstein population worldwide (VanRaden & Null 2015). The bovine *APOB* variant represents a loss of function mutation similar to *APOB*-associated familial hypobetalipoproteinemia-1, showing incomplete penetrance with respect to expression of clinical signs by a particular age of the affected people (OMIM 615558). Therefore, it was assumed that the bovine *APOB*-associated CD defect could be explained by an autosomal monogenic recessive mode of inheritance, although clinically normal *APOB* heterozygotes showed reduced cholesterol and triglyceride blood concentrations, indicating a possible codominant inheritance for cholesterol and triglyceride levels (Gross *et al.* 2016; Kipp *et al.* 2016; Saleem *et al.* 2016).

Since the identification of the causal variant in the *APOB* gene, we genotyped a total of 91 Holstein calves that appeared much smaller than their contemporaries, had owner-reported clinical signs of chronic diarrhea and illthrift, and from which blood samples were submitted for diagnostic purposes (S2). In addition, some owners of these 91 calves were suspicious of CD based on the fact that both parents were related to Maughlin Storm. The age of these

Address for correspondence

C. Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern 3001, Switzerland.
E-mail: cord.droegemueller@vetsuisse.unibe.ch

Accepted for publication 13 March 2019

91 cases varied from two to 921 days with a mean of 84.8 and a median of 47 (S1). Genomic DNA was extracted from EDTA-stabilized blood samples using the Maxwell instrument (Promega) and analyzed with the direct test targeting the *APOB* variant according to Menzi *et al.* (2016). The three possible *APOB* genotypes (CDF: non-carrier; CDC: heterozygous carrier; CDS: homozygous carrier) are presented according to standards of the World Holstein Friesian Federation (n.d.). About 80% of the submitted calves (73/91) were genotyped as homozygous mutant (CDS), confirming the initial suspicion of CD. Interestingly, 20% of clinically affected calves (18/91) were heterozygous mutant (CDC). The mean age of diagnosis of the 73 CDS affected calves was 69 days (SD ± 109.59), and for the 18 CDC clinically affected calves 148 days (SD ± 173.83), which did not differ significantly (S1).

Besides the six *APOB*-homozygous mutant CD-affected animals reported earlier (Mock *et al.* 2016), we were able to examine an additional 15 CD-suspect calves clinically (S2). Six of these 15 cases were genotyped as CDS, whereas nine were CDC. The clinical phenotype of all 15 cases included poor development, low weight and intermittent diarrhea, as reported before (Mock *et al.* 2016). Laboratory examinations for rotavirus, coronavirus, bovine viral diarrhea virus, coccidia, cryptosporidia, *E. coli* and *Salmonella* spp. gave negative results, and symptomatic treatments did not lead to clinical improvement. Three affected animals with the CDC genotype were euthanized, and the pathological phenotype of these animals was similar to what we have observed before in CDS-affected cases (Mock *et al.* 2016).

Eleven of the 12 *APOB* homozygous mutant CD-affected animals had cholesterol and triglyceride readings (S2). Mean blood total cholesterol [norm based on Mock *et al.* (2016): 1.20–3.84 mmol/l] for the 11 CDS cases was 0.21 mmol/l (SD ± 0.09), whereas the mean value for the nine CDC affected calves was 1.34 mmol/l (SD ± 0.76) (S2). The mean measured blood triglyceride concentration [norm based on Mock *et al.* (2016): 0.19–0.51 mmol/l] was 0.07 mmol/l (SD ± 0.02) for the 11 CDS cases and 0.17 mmol/l (SD ± 0.11) for the nine CDC affected calves

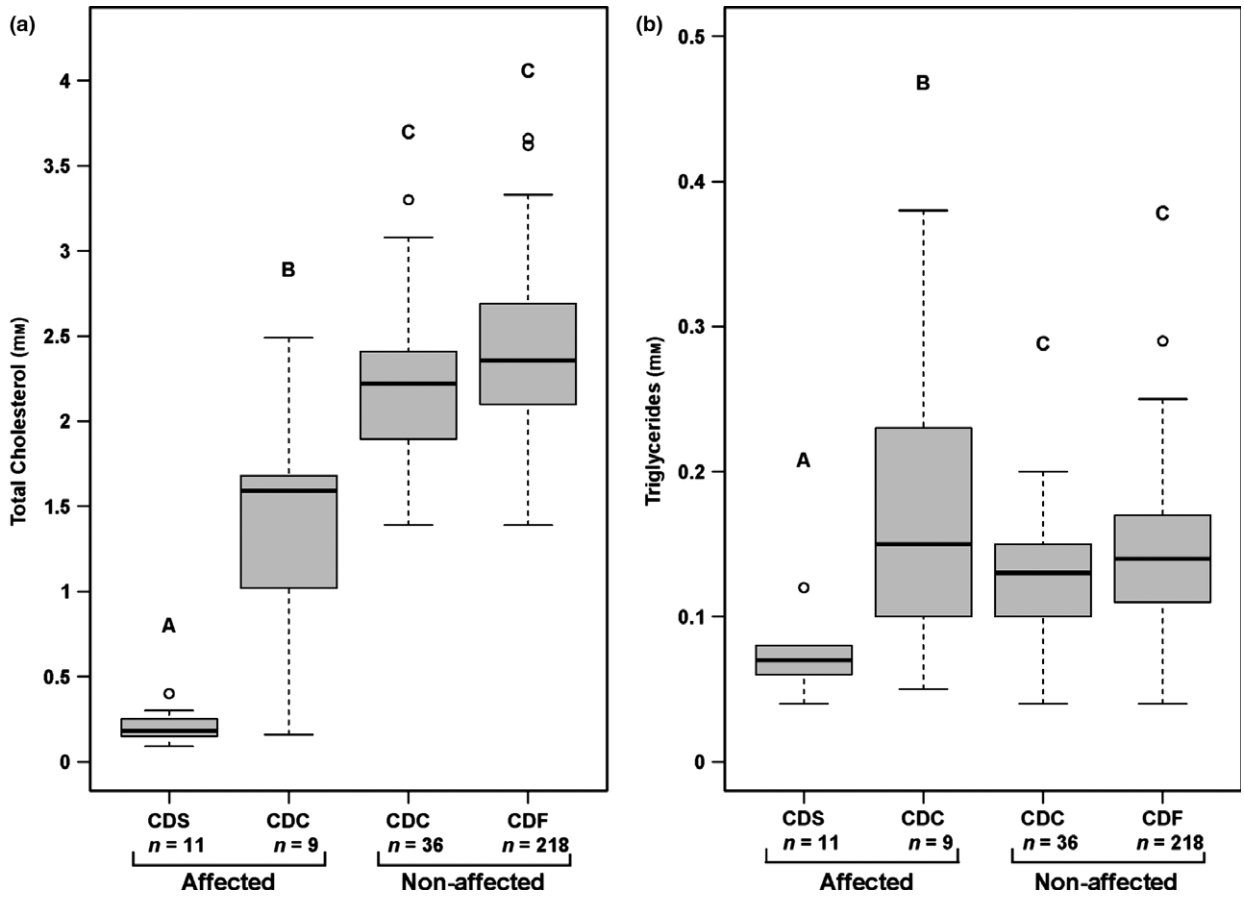


Figure 1 Boxplots showing (a) total blood cholesterol and (b) triglyceride concentrations in clinically CD affected and non-affected Holstein cattle. *APOB* genotypes are shown as wild-type (CDF), heterozygous mutant (CDC) and homozygous mutant (CDS). The number of animals per group are shown below the boxplots (11 CDS affected and nine CDC affected; 36 CDC non-affected and 218 CDC non-affected). Different capital letters above the box plots indicate significant differences between the groups ($P < 0.025$).

(S2). These values for total blood cholesterol and triglycerides were subsequently compared with similar data from adult Holstein males determined in the course of a previous study (Gross *et al*. 2016). That control cohort of 254 artificial insemination sires included 218 bulls genotyped as CDF and 36 as CDC (S2). A one-way ANOVA with total cholesterol and triglyceride concentrations as the response variables and the different *APOB* genotypes (CDS, CDC, CDF) in combination with the clinical status of the animals (CD-affected, non-affected) as the explanatory variable resulting in four groups (CDS-affected, CDC-affected, CDC-normal, CDF-normal), was conducted with the Anova function of the CAR package (Fox & Weisberg 2011). The significant outcome for total blood cholesterol ($F = 96.51$, DF = 270, $P < 2.2e\text{-}16$) and triglyceride concentrations ($F = 10.6$, DF = 270, $P = 1.311e\text{-}06$) led to the interpretation that at least both means differed significantly among *APOB* genotypes (Fig. 1). To further investigate the differences between the genotypes, a post-hoc Tukey test was performed. The function imple.glht from the MIXLM package (Liland 2018) was used for the pairwise comparisons. The pairwise comparison showed a significant difference in blood total cholesterol concentrations between the non-affected CDC controls and the CD-affected CDC cases (S3). This effect was not observed in triglyceride concentrations, as there was more variation in the triglyceride levels in the CD-affected animals than in any other group. The distribution of total cholesterol and triglyceride levels for the four groups investigated in the ANOVA is shown in Fig. 1. In comparing the four groups, a tendency of total cholesterol concentrations increasing with each additional copy of the *APOB* wild-type allele (Fig. 1) was observed, indicating an additive (codominant) effect. However, a significant difference was also observed between the two groups of animals with the CDC genotype (affected vs. non-affected; est = 0.909 with SE $\pm$ 0.166, z-value = 5.465, $P < 0.0001$) in the pairwise comparison (Fig. 1).

Based on these observations, we assume that an animal with a single copy of the *APOB* variant might show the CD disease due to disturbed lipid homeostasis. Based on our observations, most likely some of these individuals show a fatal inability to maintain a blood cholesterol level sufficient for life. On the other hand, it is known that the cholesterol level is age-dependent and that the cholesterol level in cattle levels off at about 21 days of age (Shope 1928), and the values of the nine CDC affected calves were determined much later in life (S2). Nonetheless, the CDC-clinically affected animals were noted much earlier by the owners, at a similar age as the CDS-clinically affected cases, mostly during the first weeks of life when they were fed milk (S1). Nonetheless, most heterozygous animals develop normally, and therefore the mutation acts most likely as incomplete dominance with reduced penetrance in heterozygotes. As all homozygotes develop the clinical disorder, the penetrance of the *APOB* variant is fully complete in homozygotes. On the other hand, only some heterozygotes show a clinical phenotype, meaning that penetrance is incomplete in heterozygotes. Recent findings in human genetics indicate that incomplete penetrance for presumed Mendelian diseases is likely more common than previously believed (Chen *et al*. 2016). Alternatively, possible allelic and/or genetic heterogeneity could also explain the occurrence of the CD disease in some *APOB* heterozygotes

We conclude that cholesterol deficiency not only affects *APOB* mutant homozygotes but occurs also in heterozygotes. The CD-affected heterozygous animals can show clinical signs similar to the homozygous mutants. Furthermore, we showed a significant difference in blood total cholesterol and triglyceride concentrations between clinically CD-affected and non-affected *APOB* heterozygous animals. Therefore, these findings support an incomplete dominance mode of inheritance with incomplete penetrance in heterozygotes.

Acknowledgements

The authors are grateful to Swissgenetics and the Swiss Cattle Breeding Associations (www.swissherdbook.ch and www.holstein.ch) for financial support. The authors thank J.J. Gross and R.M. Bruckmaier for providing the blood total cholesterol and triglyceride data of the control cattle. Furthermore, we thank all farmers and veterinarians for sending blood samples of suspicious cases for genotyping. Finally, we thank Nathalie Besuchet-Schmutz, Muriel Fragnière and Sabrina Schenk for expert technical assistance. This study was partially funded by the Swiss National Science Foundation.

References

Charlier C. (2016) The role of mobile genetic elements in the bovine genome, *Plant and Animal Genome Conference XXIV*, January 9–13, 2016 San Diego Abstract W636.

Chen R., Shi L., Hakenberg J. *et al*. (2016) Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nature Biotechnology* **34**, 531–8.

Duff J.P., Passant S., Wessels M., Charlier C., Hateley G. & Irvine R.M. (2016) Cholesterol deficiency causing calf illthrift and diarrhoea. *Veterinary Record* **178**, 424–5.

Fox J. & Weisberg S. (2011) *An R Companion to Applied Regression*, 2nd edn. Sage, Thousand Oaks, CA. (http://socserv.socsci.mcmaster.ca/jfox/Books/Companion).

Gross J.J., Schwinn A.-C., Schmitz-Hsu F., Menzi F., Drögemüller C., Albrecht C. & Bruckmaier R.M. (2016) Rapid communication: cholesterol deficiency–associated *APOB* mutation impacts lipid metabolism in Holstein calves and breeding bulls. *Journal of Animal Science* **94**, 1761–6.

Kipp S., Segelke D., Schierenbeck S. *et al*. (2016) Identification of a haplotype associated with cholesterol deficiency and increased juvenile mortality in Holstein cattle. *Journal of Dairy Science* **99**, 1–17.

Liland K.H. (2018) MIXLM: MIXED MODEL ANOVA AND STATISTICS FOR EDUCATION (R package version 1.2.3). (https://CRAN.R-project.org/package=mixlm)

Menzi F., Besuchet-Schmutz N., Fragnière M. *et al.* (2016) A transposable element insertion in *APOB* causes cholesterol deficiency in Holstein cattle. *Animal Genetics* **47**, 253–7.

Mock T., Mehinagic K., Menzi F., Studer E., Oevermann A., Stoffel M.H., Drögemüller C., Meylan M. & Regenscheit N. (2016) Clinicopathological phenotype of autosomal recessive cholesterol deficiency in Holstein cattle. *Journal of Veterinary Internal Medicine* **30**, 1369–75.

Saleem S., Heuer C., Sun C., Kendall D., Moreno J. & Vishwanath R. (2016) Technical note: the role of circulating low-density lipoprotein levels as a phenotypic marker for Holstein cholesterol deficiency in dairy cattle. *Journal of Dairy Science* **99**, 5545–50.

Schütz E., Wehrhahn C., Wanjek M., Bortfeld R., Wemheuer W.E., Beck J. & Brenig B. (2016) The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of *TBF1M* and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of *APOB*. *PLoS ONE* **11**, e0154602.

Shope R.E. (1928) The effect of age on the total and combined cholesterol of the blood serum. *Journal of Biological Chemistry* **80**, 141–8.

VanRaden P. & Null D. (2015) Holstein Haplotype for Cholesterol Deficiency (HCD). (https://queries.uscdcb.com/reference/changes/HCD_inheritance.pdf).

World Holstein Friesian Federation (n.d.) Standardized Labelling for Genetic Trait Coding. (http://www.whff.info/documentation/documents/WHFFStandardizeLabellingofGeneticTraitCodingv4.pdf)

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Boxplots showing the age distribution in 91 CD-affected and 254 non-affected Holstein cattle.

Table S1. List of animals used in the study.

Table S2. Output of the Tukey test showing the pairwise comparisons in detail.

3.3 Results reverse genetic approach

Mining massive genomic data of two Swiss Braunvieh cattle populations reveals six novel candidate variants that impair reproductive success

Journal:	BMC Genetics Selection Evolution
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, visualizations, writing original draft and revised versions
Displayed version:	published version
DOI:	10.1186/s12711-021-00686-3

GSE Genetics Selection Evolution

RESEARCH ARTICLE Open Access

Mining massive genomic data of two Swiss Braunvieh cattle populations reveals six novel candidate variants that impair reproductive success

Irene M. Häfliger[1*], Franz R. Seefried[2], Mirjam Spengeler[2] and Cord Drögemüller[1]

Abstract

Background: This study was carried out on the two Braunvieh populations reared in Switzerland, the dairy Brown Swiss (BS) and the dual-purpose Original Braunvieh (OB). We performed a genome-wide analysis of array data of trios (sire, dam, and offspring) from the routine genomic selection to identify candidate regions showing missing homozygosity and phenotypic associations with five fertility, ten birth, and nine growth-related traits. In addition, genome-wide single SNP regression studies based on 114,890 single nucleotide polymorphisms (SNPs) for each of the two populations were performed. Furthermore, whole-genome sequencing data of 430 cattle including 70 putative haplotype carriers were mined to identify potential candidate variants that were validated by genotyping the current population using a custom array.

Results: Using a trio-based approach, we identified 38 haplotype regions for BS and five for OB that segregated at low to moderate frequencies. For the BS population, we confirmed two known haplotypes, BH1 and BH2. Twenty-four variants that potentially explained the missing homozygosity and associated traits were detected, in addition to the previously reported *TUBD1*:p.His210Arg variant associated with BH2. For example, for BS we identified a stop-gain variant (p.Arg57*) in the *MRPL55* gene in the haplotype region on chromosome 7. This region is associated with the 'interval between first and last insemination' trait in our data, and the *MRPL55* gene is known to be associated with early pregnancy loss in mice. In addition, we discuss candidate missense variants in the *CPT1C*, *MARS2*, and *ACSL5* genes for haplotypes mapped in BS. In OB, we highlight a haplotype region on chromosome 19, which is potentially caused by a frameshift variant (p.Lys828fs) in the *LIG3* gene, which is reported to be associated with early embryonic lethality in mice. Furthermore, we propose another potential causal missense variant in the *TUBGCP5* gene for a haplotype mapped in OB.

Conclusions: We describe, for the first time, several haplotype regions that segregate at low to moderate frequencies and provide evidence of causality by trait associations in the two populations of Swiss Braunvieh. We propose a list of six protein-changing variants as potentially causing missing homozygosity. These variants need to be functionally validated and incorporated in the breeding program.

Background

Good female fertility in cattle is of high interest in agriculture, to maintain production and genetic diversity within a population [1]. Therefore, traits that affect growth, birth, and fertility are important to breeders

*Correspondence: irene.haefliger@vetsuisse.unibe.ch
[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland
Full list of author information is available at the end of the article

and semen suppliers [1]. In the current Swiss breeding programs, the traits used or intended for breeding value estimation are e.g., non-return rate, the time between calving, birth weight, calving ease, multiple birthing, rearing success, traits regarding survival within a defined period, and slaughter weight. Several studies have shown that fertility ability in dairy cattle has declined in international breeds [2, 3], the main reason being the unfavourable correlation between milk production and fertility [2–4]. This has led to increased efforts to include and prioritize fertility traits in breeding schemes and to identify the genetic causes in female reproduction, in order to breed more fertile cows. Generally, these traits have a low heritability, however, with the improvement of genomic methods, including the implementation of genomic selection (GS) [5], the reliability of breeding values can be increased [6]. Genomic selection is based on single nucleotide polymorphism (SNP) array genotyping and is especially interesting for highly complex traits with a low to moderate heritability, since it can partially account for Mendelian sampling [5, 6]. Combining comprehensive genotyping of breeding animals and recording of the phenotypes within a breeding program provides a promising database for genetic evaluations, including genome-wide association studies (GWAS). These can be carried out for all recorded phenotypes included in a breeding scheme to identify the underlying quantitative trait loci (QTL). Furthermore, the genetic variation of the population is well represented within the routinely genotyped cattle population. Studies on the search for haplotypes that show a significant reduction in homozygosity have been carried out in many species and breeds (e.g. [7–27]). Haplotypes that show no homozygous carriers may indicate the presence of recessive lethal variants within these regions. In this way, many causative variants have been detected by mining these haplotypes and using whole-genome sequencing data (e.g. [9–18, 26–32]). In the international Brown Swiss population, two haplotype regions have been repeatedly mapped, namely Brown Swiss Haplotype 1 (BH1) on chromosome 7 and Brown Swiss Haplotype 2 (BH2) on chromosome 19 [11, 19]. For BH2, which is associated with high juvenile mortality, a likely causative variant was found in the *TUBD1* gene [11, 19]. Likewise, various embryonic lethal protein-changing variants have been reported to be associated with haplotypes showing missing homozygosity in other cattle populations [10–18, 27, 33]. These examples demonstrate that short-read-based whole-genome sequencing (WGS) data make it possible not only to identify genomic regions, but also to decipher the entire genetic code of an individual in a time- and cost-efficient manner. Thereby the likely causative variants for monogenic traits segregating in livestock populations can be discovered and systematically selected against [23].

In Switzerland, Original Braunvieh (OB) and Brown Swiss (BS) are agronomically important dual purpose and dairy cattle populations, respectively. Historically, the OB population is the autochthonous brown Swiss cow with no BS influence and is adapted to the Swiss climate, whereas the modern BS originates from the OB by introgression of North American BS cattle [34]. It has been shown that the two populations segregate genetically from each other and are thus considered independent populations [35]. Both populations occur predominantly in the eastern parts of Switzerland and still have an important role in traditional alpine farming, which is frequently practiced.

The aims of this study were to identify haplotypes associated with growth, birth and fertility traits in two local Braunvieh cattle populations. Based on missing homozygosity screens that focus on trios, linkage disequilibrium analyses, and trait associations, we detected both known and novel haplotype regions segregating at low to moderate frequencies with a negative influence on female reproduction and calf survival. Based on WGS data, we identified potential protein-coding candidate causative variants for selection, including those that impair early embryonic and pre-weaning survival.

Methods

To give a comprehensive overview of the methods applied, the workflow used to mine the genomic data from SNP arrays, WGS data, and a broad variety of phenotypes for the Swiss BS and OB populations is summarized in Fig. 1.

SNP array data

The SNP array data used in this study were provided by the breeding associations of the BS and OB populations. The cleaned SNP dataset comprised 114,890 SNPs and represents a combination of different SNP arrays. As of 2018, the positions of SNPs were updated to the latest cattle reference sequence ARS-UCD1.2 [36, 37]. Data were imputed by using the software Fimpute v2.2 [38] to correct for erroneously called SNPs and to complete the SNP set for individuals genotyped with a lower chip density. Data were phased by using the Beagle v5.0 software package [39].

The OB dataset included 10,085 genotyped animals, from which 3287 trios (trio: sire, dam, and offspring) and 4360 paternal half-sib groups (pgp: **p**arent – **g**rand-**p**arent; sire, maternal grandsire, and offspring) were derived for the analyses (Table 1). The BS dataset included 48,807

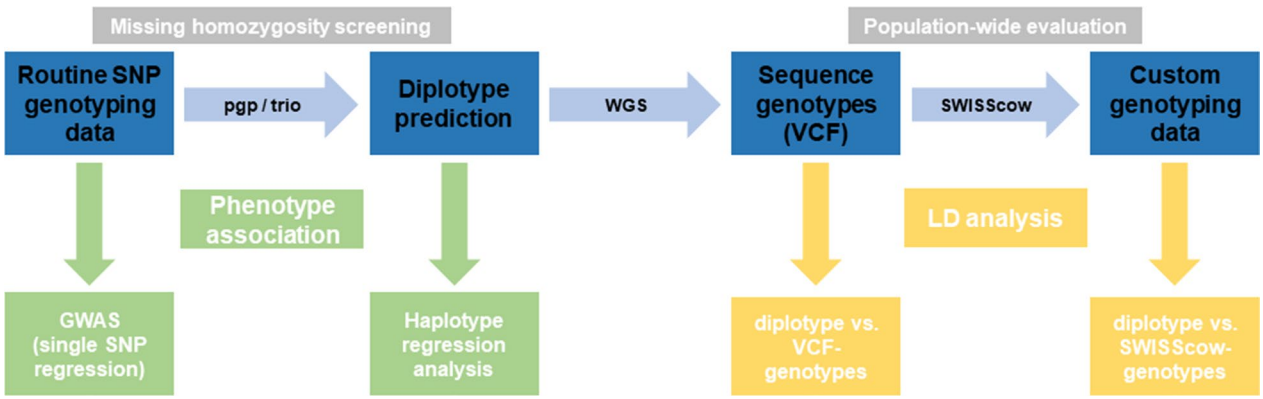


Fig. 1 Workflow of the reverse genetic analyses. Note that the steps of the main genomic analyses that are described in the Methods section are indicated in blue. Further evaluations are shown in green (phenotypic analysis) and yellow (linkage disequilibrium (LD) analysis)

Table 1 Number of genotyped animals from the two Braunvieh cattle populations used for missing homozygosity scan, haplotype, and GWAS analyses

	Original Braunvieh (OB)	Brown Swiss (BS)
Number routine genotyped animals	10,085	48,807
Number trios (sire, dam and offspring)	3287	14,450
Number paternal half sib groups—pgp (sire, maternal grandsire and offspring)	4360	32,319

genotyped animals, with 14,450 trios and 32,319 half-sib groups (Table 1). The SNP array data were quality filtered by retaining SNPs with a MAF higher than 0.01 and a call rate greater than 0.9, and individuals with a call rate greater than 0.8.

Haplotype analysis

We used the snp1101 software to screen the genomic data for haplotypes that showed a significant deviation from the Hardy–Weinberg equilibrium (HWE) based on the exact test of HWE [40, 41]. The screen was performed for each population and repeated for both the trio-based and the pgp-based approaches. The dataset used for screening was generated by including the genotypes of the Swiss animals born after 2009 and their ancestors. The snp1101 software screens each autosome based on a sliding-window approach, for which we defined a window size of 50 SNPs that represents genome regions ranging from 0.26 to 3.2 Mb with an average length of 1.1 Mb. The output is a list of all the significant haplotypes with the number of observed and expected homozygous carriers, the allele frequency and the haplotypes themselves. We carefully defined all the regions of interest as the regions that show a reduction in homozygosity and a significant deviation from HWE with their adjusted p-value for a false discovery rate (FDR) corrected according to the Benjamini-Yekutieli procedure [42]. For each region, we chose the most significant haplotype to reduce the number of false positive carriers. We defined the diplotype status for each haplotype of all the animals in the population to select the most relevant animals for whole-genome sequencing. By sequencing healthy breeding animals that are either heterozygous or homozygous carriers of the diplotypes, we increased the probability of sequencing animals that carry the associated causal variants.

Association studies

Based on the predicted diplotype status, haplotype association studies were carried out for a variety of traits and for each of the observed haplotypes. For the association analyses, routinely available conventional, non-genomic breeding values related to fertility, birth, and growth were used. The individual traits are described in Table 2. Only breeding values with a reliability higher than 0.5 were deregressed according to Garrick et al. [43] and used in a linear mixed model using the genome-wide complex trait analysis (GCTA) software [44]. Thereby, only the animals that have own performance or progeny performances were included in the analysis. The model is as follows: $y_{EBV} = \mu + G + \mathbf{H\beta} + \epsilon$, where y_{EBV} is the deregressed breeding value (drEBV), μ is the average of the deregressed breeding values, G is a random genetic effect based on a genomic relationship matrix (GRM) estimated by the GCTA software [44], $\mathbf{H}$ is a vector of the diplotype status per animal, $\boldsymbol{\beta}$ is a vector of the estimated fixed effects of each haplotype and ϵ accounts for random variation.

We also performed genome-wide association studies (GWAS) on the same genomic and phenotypic data used above and we calculated the association of each single

Table 2 The 24 female fertility traits analysed

Trait group	Trait sub-group	Trait	Description
Fertility traits	Fertility	Non-return rate heifer (nrr)	Heifers non-return rate after 56 days, binary
		Non-return rate cow (nrk)	Cows non-return rate after 56 days, binary
		Interval first to last insemination heifer (vzr)	Interval between first and last insemination for heifer, days
		Interval first to last insemination cow (vzk)	Interval between first and last insemination for cows, days
		Interval calving to insemination (raz)	Interval from calving to first service, days
Birth traits	Birth history direct	Percentage normal births (ngd)	Calving ease, scored between 1-without help to 5-dystocia
		Percentage live births (lgd)	Percentage of calves born alive
		Birth weight (ggd)	Weight of calve at birth, kg
		Gestation length (tdd)	Days from successful insemination to birth
		Multiple birth (twind)	Percentage of multiple births
	Birth history maternal	Percentage normal births (ngm)	Calving ease, scored between 1-without help to 5-dystocia
		Percentage live births (lgm)	Percentage of calves born alive
		Birth weight (ggm)	Weight of calve at birth, kg
		Gestation length (tdm)	Days from successful insemination to birth
		Multiple birth (twinm)	Percentage of multiple births
Growth-related traits	Rearing success	Survival period 1 (p1)	Survival from day 3 up to day 30
		Survival heifer period 2 (hp2)	Survival of heifers from day 31 up to 458 days
		Survival bull period 2 (bp2)	Survival of young bulls from 31 days up to 183 days
	Slaughter traits calves	Slaughter weight (cwco)	Weight at slaughter, kg
		Carcass conformation score (ccco)	Amount of meat at slaughter, kg
		Carcass fat score (cfco)	Fat cover in the meat
	Slaughter traits adults	Slaughter weight (cwao)	Weight at slaughter, kg
		Carcass conformation score (ccao)	Amount of meat at slaughter, kg
		Carcass fat score (cfao)	Fat cover in the meat

SNP to the deregressed breeding value using the model: $y_{EBV} = \mu + G + S\boldsymbol{\beta} + \epsilon$, where y_{EBV} is the drEBV, μ is the average of the drEBV, G is a random effect based on a GRM estimated by the method in [45], S is a scalar of the genotype status of each SNP per animal, $\boldsymbol{\beta}$ is the estimated fixed effect of each SNP and ϵ accounts for random variation. For this analysis, we used the snp1101 software [40] and applied a Bonferroni correction for multiple testing with an increased significance level at $p < 4.35e{-}7$. All the genotyped animals were included in these models. The GWAS results were visualized by creating Manhattan plots with the qqman package [46].

Preparation of the whole-genome sequencing data

To identify the putative causative variants, we performed whole-genome sequencing on selected animals that were predicted to be carriers of the identified haplotypes. For this purpose, we sequenced 70 animals in addition to the 360 genomes that had been sequenced within other projects or which are publicly available, to reach a total of 430 genomes of healthy breeding animals (see Additional file 1: Table S1). Among these, 114 and 100 were from purebred BS and OB animals, respectively. First, the reads were trimmed with fastp [47] by applying the following quality thresholds with the flags: --qualified_quality_phred 20, --length_required 35, --cut_window_size 3, --cut_mean_quality 15, --cut_by_quality5 20, --cut_by_quality3 20, and for the WGS data from the NovaSeq6000, the flag --trim_poly_g was added. The quality-controlled reads were mapped to the latest cattle reference sequence ARS-UCD1.2 [36, 37] using the Burrows-Wheeler Aligner v. 0.7.17 [48], deduplicated by using the MarkDuplicates tool from the Picard software v.2.18.2 [49], recalibrated using the BaseRecalibrator and PrintReads applications of the GATK v3.8.1.0.gf15c1c3ef software [50] with known variants from the 1000 Bull Genomes Project run 7 (BQSR file version 2) [51], and sorted by using the samtools v1.8 software [52].

The average read depth and insert size were calculated using covstats in the goleft v0.1.19 tool [53]; read depth ranged from 6.8 × to 72.8×, with an average of 18.6x. For genotyping, the HaplotypeCalling, CombineGVCFs, CatVariants and GenotypeGVCFs tools from GATK v3.8.1.0.gf15c1c3ef were used to produce a variant call format (VCF) file [50]. To predict the effects of the detected variants, we used the NCBI Annotation Release 106 [54] and the SnpEff v4.3 software [55]. Furthermore, quality scores were estimated for each variant

by applying the GATK recommendation for hard-filtering of indels and small nucleotide variants (SNV) with the SelectVariants and VariantFiltration tools of GATK v3.8.1.0.gf15c1c3ef [50]. This workflow is consistent with the workflow proposed by the 1000 Bull Genomes Project (run 7) [51, 56].

Analysis of the WGS data for the design of the custom array

An algorithm was implemented to efficiently screen the identified haplotype regions with a significant depletion in homozygous animals for candidate variants. Thus, we increased the window size by 2 million bp on each side and we looked for variants that did not occur in the homozygous state in any of the 430 animals. As the read depth varied, we allowed for missing values (a maximum of 50% of genotypes) but assumed that all variants passed the quality score criteria. In addition, we postulated that at least one animal in each population, BS or OB, had to be a carrier of potential variants, but that not more than 75% of the animals should be carriers.

Subsequently, we designed a custom Affymetrix array called "SWISScow chip" in cooperation with the Swiss breeding associations. We selected 465,768 variants to design this custom array, of which 236,043 were selected from the haplotype regions mapped in this study and 230,305 were protein-coding variants spread genome-wide. The design process allowed us to consider 44% of the initially selected variants. The final SNP panel encompassed 318,216 variants, including 112,854 "routine" variants that were genotyped before using different Illumina bovine beadchips plus 205,362 project-specific variants, so called "research" markers.

This array was used in the Swiss routine genotyping throughout 2020, and provided us with the genomic information of 13,667 animals including 6575 BS and 1489 OB plus a cohort of 5603 Swiss Holstein (HO) cattle.

Linkage disequilibrium (LD) analyses

To confirm the association between the variants in the VCF file and the defined haplotypes, we calculated the linkage disequilibrium between the defined haplotypes' diplotypes and the variants in the VCF file, and with the genotypes from the custom array. While linkage analysis of the VCF file included 79 BS and 94 OB samples, further analysis depended on the quality control performed on the VCF data of the chromosome of interest (plink function --mind 0.1) [57]. This information was integrated into the data from the custom array, including the 8064 Braunvieh genotypes. Linkage analyses were performed using the plink v1.9 software [57].

Visualization of the results

To provide a comprehensive overview of the results, the OmicCircos Rpackage was used to visualize the identified genomic regions [58]. Figures 2 and 3 were obtained by including the genomic regions of all significant haplotypes from the trio approach and, if in the same haplotype region, a significant haplotype with the pgp approach was found, this is indicated in the plots. Furthermore, the plots include the LD results between markers on the custom array with the diplotypes, the significant haplotype association for the various traits, and the significant GWAS association combined in the above-defined trait groups (Table 2) for BS and OB, respectively.

Interpretation of the variants

For the interpretation of the genomic variants, the UCSC Genome Browser tool suite was used to calculate conservation scores [59]. First, the LiftOver UCSC tool and their LiftOver file for BosTau9 to HG38, which is the DNA base conversion from the bovine reference sequence ARS-UCD1.2 [36] to the human genome 38 [60], was applied by command-line coordinate, lifting all the variants from the VCF file. Furthermore, for these new positions, the base-wise conservation scores of 99 vertebrate genomes with the human genome were extracted from the UCSC database and resulted in the phyloP and PhastCons scores shown in Table 3 and Table S5 (see Additional file 6: Table S5) [61, 62]. Both scores represent a comparative genomic alignment approach, but they use different algorithms. PhyloP measures the conservation and evolutionary acceleration at individual sites with absolute values of −log(p-value), and PhastCons identifies stretches of conservation with values demonstrating probabilities of negative selection from 0 to 1 [61, 62].

For the variants that show a linkage between the WGS data and the haplotypes in relevant regions, we extracted information about their effects by manually using the web-based prediction tools PROVEAN [63] and MutPred2 [64] resulting in the SiftScores and MutPred2 scores shown in Table S5 (see Additional file 6: Table S5).

Finally, we analysed the segregation of candidate variants within the international cattle population of the 1000 Bull Genomes project run 8 that includes 4109 animals from various breeds [56].

Results

Detection of numerous novel haplotypes associated with fertility, birth, and growth-related traits

We identified many genomic candidate regions that show a significant depletion in homozygosity, significant haplotype association, and a genome-wide association to fertility, birth, and growth-related traits. For the

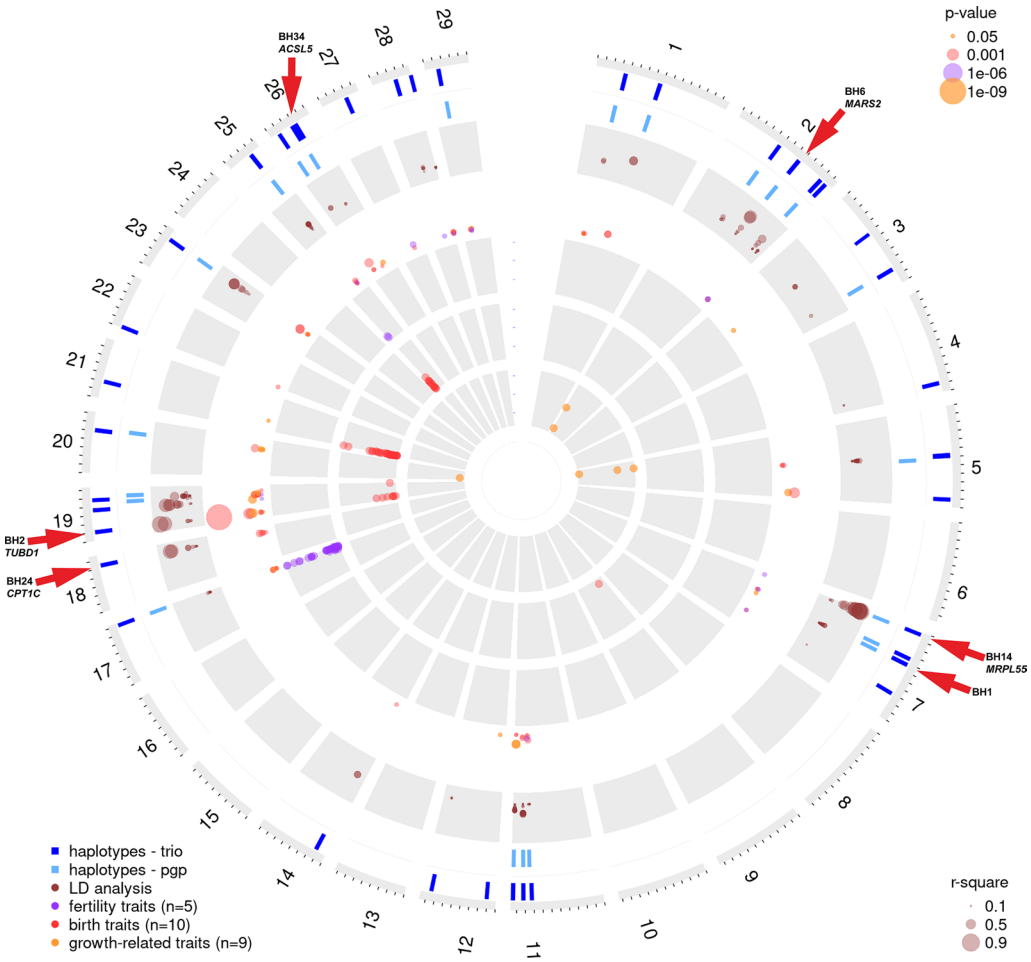


Fig. 2 Genome-wide summary of the data mining for the Brown Swiss population. In the outer circles, the identified haplotypes per chromosome with reduced homozygosity for the trio-based approach are indicated in dark blue and their associated haplotypes of the pgp-based approach in light blue. Note that only the haplotypes that were detected through the pgp-based approach are shown if, within the same region, another haplotype was detected by the trio-based approach. The circle with the brown dots indicates LD (r^2) between haplotypes and markers on the custom SNP array. Note that the dot size correlates with the extent of LD. The third circle shows the significant haplotype association results. Note that the different colors represent the three groups of evaluated traits and the dot size correlates with the significance values accordingly. The three inner circles present the significant GWAS results across all fertility (purple), birth (red), and growth-related (yellow) traits. Scales are based on the −log10(p-value). Note that the red arrows indicate the previously identified haplotypes BH1 and BH2 [19] and the herein described *MRPL55*-related haplotype BH14, as well as the BH14, BH24 and BH34 haplotypes and their associated genes that harbor the most likely causative variants

two Braunvieh populations studied here, BS and OB, each of the haplotype analyses based on the pgp- and the trio-based approaches resulted in a concatenated list of all the haplotypes that deviated significantly from HWE. The identified haplotypes were named according to previous studies [19] as BH haplotypes for BS and OH haplotypes for OB (Tables 3 and 4). For the BS population, we detected 38 haplotypes with the trio-based approach (Table 3) and 53 haplotypes with the pgp-based approach, with 19 overlapping haplotypes (Fig. 2) and (see Additional file 2: Table S2). Two of the identified haplotypes, BH1 on chromosome 7 and BH2 on chromosome 19 had already been detected in 2011 [19] (Fig. 2). The haplotypes that deviated most significantly from HWE were BH6, BH14, and BH1, with no homozygous carriers although at least 44 were expected (see Additional file 2: Table S2). For the OB population, we detected five significant novel haplotypes with the trio-based approach (OH2 to OH6, see Table 4) and three haplotypes with the pgp-based approach, which all overlapped with those detected by the trio-based approach (Fig. 3) and (see Additional file 2: Table S2). The haplotypes that deviated most significantly from HWE were OH2 and OH4, with no observed homozygous diplotype

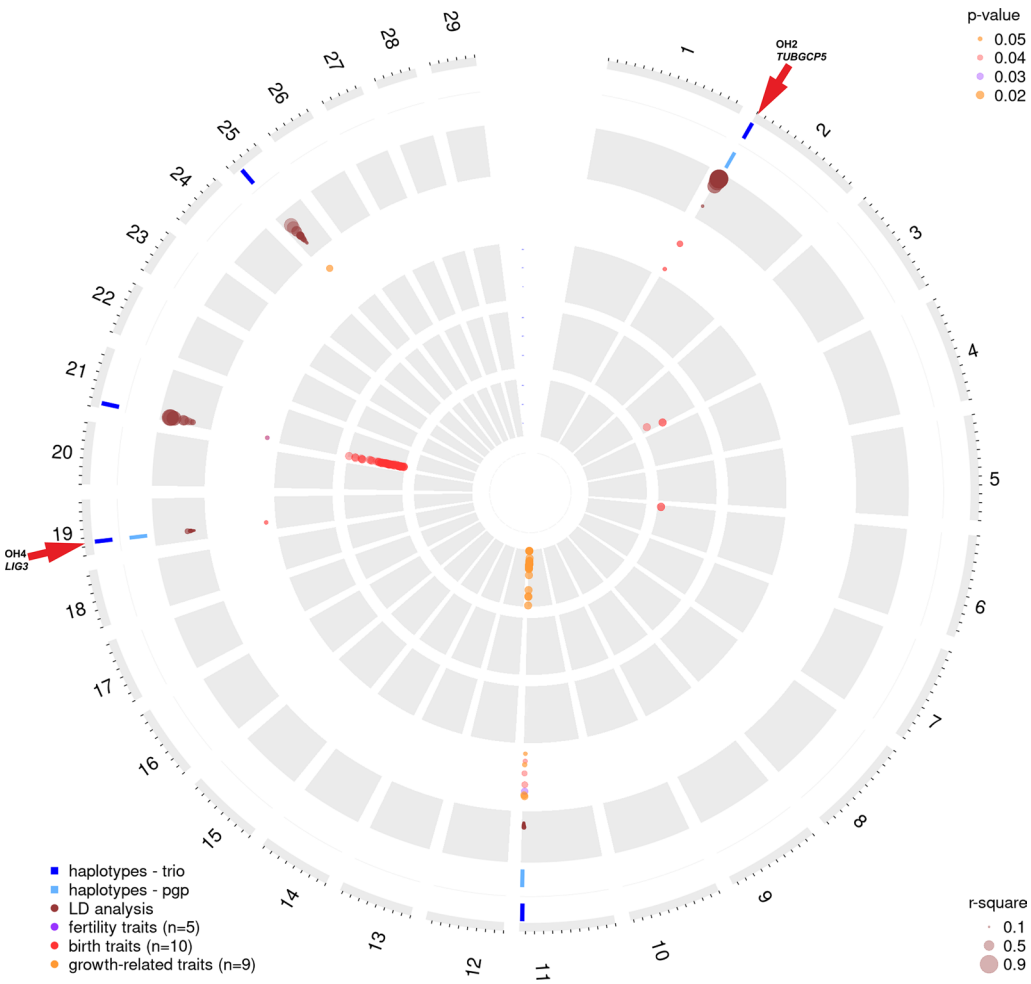


Fig. 3 Genome-wide summary of the data mining for the Original Braunvieh population. In the outer circles, the identified haplotypes per chromosome with reduced homozygosity for the trio-based approach are indicated in dark blue and their associated haplotypes of the pgp-based approach in light blue. Note that only the haplotypes that were detected through the pgp-based approach are shown if, within the same region, another haplotype was detected by the trio-based approach. The circle with the brown dots indicates LD (r^2) between haplotypes and markers on the custom SNP array. Note that the dot size correlates with the extent of LD. The third circle shows the significant haplotype association results. Note that the different colors represent the three groups of evaluated traits and the dot size correlates with the significance values accordingly. The three inner circles present the significant GWAS results across all fertility (purple), birth (red), and growth-related (yellow) traits. Scales are based on the −log10(p-value). Note that the red arrows indicate the *LIG3*-related haplotype OH4 described in this paper, as well as the OH2 haplotype and its associated gene *TUBGCP5*

carriers, although at least 15 were expected (see Additional file 2: Table S2). After correction for FDR, three interesting haplotype regions (OH7, OH8 and OH9) with a non-significant depletion of homozygosity were found for OB (see Additional file 2: Table S2; Additional file 3: Table S3; and Additional file 5: Table S4).

The complete GWAS results for the 24 traits, conducted on the identified BH and OH haplotypes are in Table S3 (see Additional file 3: Table S3). BH and OH haplotypes showed significant associations on 21 and five different chromosomes, respectively (Figs. 2 and 3). The strongest significant association was found for BH2 with four birth and two growth-related traits (see Additional file 3: Table S3), which confirms the previous finding that the BH2-related missense variant p.His210Arg in the *TUBD1* gene causes ill-thrift and juvenile mortality in Brown Swiss cattle [11].

All the results of the single SNP GWAS are visualized in traditional Manhattan plots (see Additional file 4: Fig. S1) and the most significant markers were selected for a comprehensive overview per trait group (Figs. 2 and 3) and (see Additional file 5: Table S4). For example, obvious GWAS hits for fertility traits in BS, such as non-return rate and interval first to last insemination

Table 3 List of haplotypes in the Brown Swiss population identified by the trio-based approach

Name	Chr	Start	End	Allele frequency	Proposed associated gene[a]
BH1[b]	7	41371808	42545291	3.02	*TCF3*
BH2[c]	19	9726237	10819756	4.22	*TUBD1*
BH3	1	34908448	36034631	2.05	
BH4	1	76587512	77861430	2.20	*GMNC*
BH5	2	57787148	58768161	2.87	*LRP1B*
BH6	2	86065338	87460373	3.41	*MARS2*
BH7	2	120284002	121473182	1.93	*SPATA3*
BH8	2	127933873	129004898	2.98	
BH9	3	52432936	52,689742	1.43	*HFM1*
BH10	3	101636810	102827915	2.26	
BH11	4	108811998	109953666	2.66	
BH12	5	57653479	59777422	2.32	*RNF41* and *LRP1*
BH13	5	110617654	111664596	1.63	
BH14	7	2588873	3357718	3.04	*MRPL55*
BH15	7	33954389	34956540	4.12	
BH16	7	79845977	80708046	2.20	
BH17	11	82034654	82881173	2.39	*PGGHG*
BH18	11	91766016	92804942	2.54	
BH19	11	103891372	105003404	1.77	
BH20	12	10254014	10967287	2.02	
BH21	12	74954939	75817277	1.61	
BH22	14	12290759	13350202	2.64	
BH23	17	69786512	70645226	2.98	
BH24	18	52977041	54113281	2.32	*CPT1C*
BH25	19	35808547	36844178	1.86	*SNF8*
BH26	19	47874708	48724082	1.89	
BH27	20	50058036	51370414	2.56	
BH28	21	19368258	20362870	1.60	
BH29	22	311724	1277064	2.07	
BH30	23	39600366	40609442	2.49	
BH31	25	27713801	29132187	1.80	*ITGAD* and *SPN*
BH32	26	9215188	10180467	1.72	
BH33	26	26659976	27689471	2.72	
BH34	26	31353340	32429589	2.37	*ACSL5*
BH35	27	28524504	29550764	1.80	
BH36	28	26661955	27743768	1.75	
BH37	28	45215132	45913154	2.07	
BH38	29	14497752	15554643	3.08	

BH BS haplotype, *Chr* chromosome number, *Start and End* start and end positions according to the reference sequence ARS-UCD1.2 [36]

[a] According to NCBI Annotation Release 106 [54]

[b] Previously described haplotype [19]

[e] Previously described haplotype and associated gene [11, 19]

Table 4 List of haplotypes in the Original Braunvieh population identified by the trio-based approach

Name[a]	Chr	Start	End	Allele frequency	Proposed associated gene[b]
OH2	2	1005580	1614673	4.48	*TUBGCP5*
OH3	11	10,406494	104418358	5.76	*MYMK*
OH4	19	14336760	15222429	3.87	*LIG3*
OH5	21	5195518	6367707	4.56	*LYSMD4*
OH6	25	9596610	10624288	3.87	*USP7*

OH OB haplotype, *Chr* chromosome number, *Start and End* start and end positions according to the reference sequence ARS-UCD1.2 [36]

[a] Since the name OH1 is reserved for a pathogenic *CNGB3*-variant [101], numbering starts at OH2

[b] According to NCBI Annotation Release 106 [54]

are on chromosome 17 where BH23 was detected, and very strong GWAS signals for birth traits in OB, such as calving ease, birth weight, and gestation length, are on chromosome 21, where OH5 was found. Among the studied growth-related traits, we identified a highly significant association signal on chromosome 11 in OB, which affects carcass conformation and co-localizes with the OH3 haplotype.

Potential candidate causative variants in *MRPL55* and *LIG3*

Subsequently, we mined the genome-wide sequence data to identify potential causative candidate variants responsible for either prenatal lethality or postnatal lethality, or sub-lethal phenotypes. Analysis of the variants that show depletion in homozygosity and are in LD with the haplotype identified protein-changing variants that are potentially causative in five BH and two OH haplotypes (Tables 5 and 6). Genotyping based on the SWISScow custom array allowed to validate whether three of the BS variants were present in several thousands of young animals. As for the haplotypes, we detected animals with homozygous variant genotypes in the current population, which lead us to anticipate either that late-onset undesired traits might occur or that these variants have incomplete penetrant effects. The disorders known to be associated with the affected genes range from embryonic lethality (e.g., *MRPL55* [65] and *LIG3* [66]) and metabolic diseases (e.g. *acyl-CoA synthetase long chain family member 5* (*ACSL5)* [67]) to neurodevelopmental disorders (e.g., *methionyl-tRNA synthetase 2* (*MARS2*) [68]) (Tables 5 and 6).

We specifically selected one variant on chromosome 7 at ~ 3 Mb in the BS population that is predicted to be deleterious due to a stop-gain variant in the *MRPL55* gene. This mutation introduces a premature stop codon (p.Arg57*), which truncates the encoded protein by almost 80% (Table 5). This variant is in perfect LD ($r^2 = 1$) with the BH14 haplotype in the WGS animals and in high LD ($r^2 = 0.9$) in the custom array genotyped animals. It

Table 5 Short list of five potential candidate causative variants for the Brown Swiss population[a]

Haplotype-region information		Gene	OMIM/OMIA	Associated disorder/gene function	Variant designation[b]			
Name	Approach				Genomic position	Transcript	Coding DNA change	Protein change
BH2	Trio	*TUBD1*[c]	607344/001939-9913	Juvenile mortality	chr19:10833921	NM_001075470.2	c.629A>G	p.His210Arg
BH6	Trio and pgp	*MARS2*	609728	Spastic ataxia 3 (lethal)	chr2:86191230	NM_001098971.1	c.1553G>A	p.Arg518Gln
BH14	Trio and pgp	*MRPL55*	611859	Early pregnancy loss	chr7:2996436	NM_001303490.1	c.169C>T	p.Arg57*
BH24	Trio	*CPT1C*	608846	Spastic paraplegia (lethal)	chr18:56098048	XM_002695120.5	c.158G>A	p.Gly53Asp
BH34	Trio	*ACSL5*	605677/002226-9615	Lipid malabsorption	chr26:32940521	NM_001075650.1	c.528C>G	p.Asn176Lys

Chr chromosome, *BH* BS haplotype

[a] Comprehensive list available in Table S5 (see Additional file 6: Table S5)

[b] According to the reference sequence ARS-UCD1.2 [36] and NCBI Annotation Release 106 [54]

[c] Previously described variant [11]

Table 6 Short list of two potential candidate causative variants for the Original Braunvieh population[a]

Haplotype-region information		Gene	OMIM/OMIA	Associated disorder/gene function	Variant designation[b]			
Name	Approach				Genomic position	Transcript	Coding DNA change	Protein change
OH2	Trio and pgp	*TUBGCP5*	608147	Proper formation of the mitotic spindles	chr2:1268426	NM_001102495.1	c.311C>A	p.Thr104Lys
OH4	Trio and pgp	*LIG3*	600940	Embryonic lethality	chr19:15080335	NM_001038107.2	c.2483_2484+4delAGGTGC	p.Lys828fs

Chr chromosome, *OH* OB haplotype

[a] Comprehensive list available in Table S5 (see Additional file 6: Table S5)

[b] According to the reference sequence ARS-UCD1.2 [36] and NCBI Annotation Release 106 [54]

was never observed in the homozygous state in any of the BS individuals. Furthermore, BH14 shows a significant negative association with the fertility trait 'interval first to last insemination' in cows and a mentionable but non-significant negative association with the fertility trait 'interval calving to first insemination' and the growth-related trait 'survival of young bulls' (see Additional file 3: Table S3). Evaluation of the *MRPL55*-related variant with the custom array revealed that it segregates within the BS population at a frequency of 0.032 without any homozygous carriers and deviates significantly from HWE (see Additional file 6: Table S5). In the OB population with more than 1400 animals genotyped with the SWISScow custom array, genotyping found only two heterozygous animals for this variant. In addition, none of the 4109 animals from various breeds within the run 8 data of the 1000 Bull Genomes project were homozygous for this variant [51].

In the BS population, we identified three additional missense variants in the *carnitine palmitoyltransferase 1C* (*CPT1C*), *MARS2*, and *ACSL5* genes (Table 5), which affect highly conserved nucleotides. Their effects on protein level were predicted to be deleterious by PROVEAN and pathogenic by MutPred2 (see Additional file 6: Table S5). Unfortunately, only one of these variants was included in the SWISScow custom array, which was associated with BH24 and located in the *CPT1C* gene (p.Gly53Asp). This variant was never detected in the homozygous state in the WGS data, and only a few homozygous carrier animals were detected in the SWISScow genotypes (see Additional file 6: Table S5). The BH24 haplotype is associated with several birth traits such as birth weight, percentage of normal births, and percentage of live births. The p.Arg518Gln variant in the *MARS2* gene that is associated with the BH6 haplotype is also of interest since this haplotype was never detected in

the homozygous state. Unfortunately, this variant was not considered in subsequent array-based genotyping. Lastly, the p.Asn176Lys variant in the *ACSL5* gene that is associated with the BH34 haplotype was also never detected in the homozygous state within the available WGS data and, to date, has not been genotyped in other animals. BH34 was rarely observed in the homozygous state in the BS population and is statistically negatively associated with the growth-related traits 'heifer survival after 30 days' and 'carcass fat score in calves'.

In the OB population, we draw attention to the haplotype region OH4 located on chromosome 19 at ~15 Mb. For this haplotype, we identified a frameshift variant (p.Lys828fs) in the *LIG3* gene leading to the change of 90 residues and premature termination of translation compared to the wild-type protein sequence, which is 26 residues longer (Table 6). This variant is in perfect LD ($r^2 = 1$) with the OH4 haplotype in the WGS animals. OH4 is negatively associated with the percentage of live births. Unfortunately, this frameshift variant was not part of the SWISScow custom array. Nevertheless, the *LIG3* variant segregates at an allele frequency of 0.035 in the WGS data of Braunvieh animals. In addition, none of the 4109 animals from various breeds within the run 8 data of the 1000 Bull Genomes project [51] were homozygous for this variant. According to OMIM #600940, the *LIG3* gene is essential for DNA repair of mitochondrial DNA.

A missense variant that affects an evolutionary highly conserved residue (p.Thr104Lys) in the *tubulin gamma complex associated protein 5* (*TUBGCP5*) gene was detected in the OB population but was never observed in the homozygous state in the WGS data (see Additional file 6: Table S5). The associated OH2 haplotype has a negative effect on multiple births and a positive effect on the percentage of normal births. Although absent from the SWISScow array, this variant segregates at an allele frequency of 0.021 in the WGS data and is in perfect LD with the haplotype (see Additional file 6: Table S5).

In addition, for 13 other BH and two other OH haplotypes, we propose candidate causative variants (see Additional file 6: Table S5), but most of them are predicted to have a neutral or benign effect or represent synonymous variants.

Discussion

This comprehensive study explored the genomic data of the two current Swiss Braunvieh dairy populations BS and OB, for reduced homozygosity due to hidden recessive variants. Such variants which, for the most part, change the amino acid sequence of proteins, lead to natural or artificial selection against homozygous individuals. This phenomenon could be due to embryonic lethality, reduced rearing success, or exclusion from the breeding population due to poor development. To better understand the functional role of the associated haplotype regions, we performed genome-wide missing homozygosity scans and subsequent comprehensive association analyses.

In the recent past, various studies that aimed at pinpointing such haplotype regions used routine SNP genotyping data and the standard approach (pgp) in which the paternal grandfather replaces the usually non-genotyped dam [19, 21, 27]. Here, we chose a trio-based approach, assuming that it would increase the power of the statistical analysis of haplotypes due to the direct relationships of the families analyzed [21]. The trio-based approach resembles the transmission ratio distortion approach used to search for lethal alleles, however, the latter is based on Bayesian statistics, while we applied the Fisher exact test of HWE [7, 8, 24, 41]. Especially for the OB population for which the amount of available data is limited, the trio-based approach is notably more efficient than the standard analysis, since more regions with lower haplotype frequencies were detected. In contrast, for the BS population, the pgp-based approach, which considers more than twice the number of family groups, detected more regions than the trio-based analysis. However, it should be kept in mind that an increased number of detected haplotype regions might include more false positives, which is supported by our results with some of the haplotype regions detected in the BS population not showing any phenotypic associations. Therefore, we focused on the haplotype regions that were detected by the trio-based approach and showed phenotypic associations with the 24 analyzed fertility, birth, and growth-related traits.

To validate the chosen approach, we searched for previously identified haplotypes and known causal recessive variants and confirmed the BH1 and BH2 haplotypes in the studied BS population [11, 19]. BH2 is known to be negatively associated with fertility and growth-related traits. We also confirmed that the variant in the *TUBD1* gene showed the highest level of LD with BH2. Interestingly, although breeding against this variant has been practiced for several years, we were still able to detect this haplotype region and its effects on the studied traits. However, strong selection of the bulls has successfully eliminated the two inherited disorders spinal muscular atrophy [69, 70] and weaver syndrome [14, 31, 71] since we did not detect any carriers. Regarding the OB population, it is known that the Fleckvieh haplotype 2 (FH2) and the associated frameshift variant in the *SLC2A2* gene segregate in the OB population at an allele frequency of 0.05 and cause the Fanconi-Bickel syndrome with growth retardation [9, 72]. Nonetheless, we did not detect the region surrounding this variant on chromosome 1. The

reason why we could not find a significant depletion in homozygosity for FH2 was that the genotyping data included several homozygous carriers of the variant associated with the Fanconi-Bickel syndrome. Due to the fact that this disorder is non-lethal, but induces liver and kidney defects leading to reduced growth [9], apparently normally developed new-born calves are genotyped before the manifestation of growth retardation.

Our aim was to detect novel causative variants, but first we confirmed the detection of the BH1 haplotype region previously reported in BS cattle [19, 33]. We found evidence for depletion of homozygotes, but a positive association with the growth-related trait carcass fat score and no indication of a negative effect on fertility traits as previously described [33]. We suggest that it is caused by the p.Asp70Asn missense variant that affects a highly conserved nucleotide in the *transcription factor 3* (*TCF3)* gene, also known as *E2A*. In men, mutations in this gene are associated with agammaglobulinemia 8 (OMIM #616941) (see Additional file 6: Table S5). *Tcf3* knock-out mice show growth retardation and increased neonatal death [73]. TCF3 is associated with the Wnt signaling pathway and thereby influences osteogenesis and is essential for B cell differentiation [73, 74]. This same variant was previously reported to be associated with BH1, but with inconclusive results [33]. In our study, the estimated LD between BH1 and this variant is rather low and the haplotype associations are merely suggestive. Therefore, either this variant causes retarded growth and increased pre-weaning lethality, but with incomplete penetrance, or is irrelevant as not all loss-of-function variants have a pathogenic effect [75]. We found several dozens of homozygous carriers of the *TCF3* variant, although it segregates with an allele frequency of 0.11 and deviates significantly from HWE. Therefore, we recommend to keep track of this variant, e.g. by clinically examining live homozygous individuals.

Regarding our extensive list of newly identified haplotype regions in the Swiss Braunvieh populations studied here, we propose six candidate causative variants (Tables 5 and 6), which are all supported by high LD values with the corresponding haplotype, and 17 variants of interest (see Additional file 6: Table S5). In the future, for many other mapped haplotypes, the detection of such variants should be done with the same approach based on population-wide evaluation, because unfortunately only every second variant is technically designable for the custom array. Furthermore, it is disputable if instead of array-based genotyping with its known limitations, one could consider genotyping by sequencing by using low-pass sequencing [76, 77]. Nonetheless, this recent method has also strong limitations such as a restricted probability to see the variant of interest on the sequence of a given individual as well as the need to be followed by imputation that is not accurate for recent variants due to the existence of ancestral and derived versions of identical haplotypes. Due to the large evidence gained by the analysis of massive phenotypic and genomic data, we recommend further evaluation of such haplotype regions and candidate variants in the future.

Among the variants suggested to cause reduced omozygosity in the BS population, we highlighted the *MRPL55* nonsense variant on bovine chromosome 7 associated with BH14. Statistically, this variant fits perfectly with the expectations for embryonic lethal variants, since LD is very high and there are no homozygous animals. The *MRPL55* gene belongs to the mitochondrial ribosomal protein (MRP) gene family, more precisely to the MRPL group, which are genes that make up the large subunit of the mitoribosome complex [78]. The mitoribosome complex is a multi-gene complex of diverse genes, denoted MRP genes [65]. Mouse knock-out experiments have revealed that different genes of the MRP family cause early embryonic death at the stage of pre-gastrulation, and expression analyses of MRP genes have shown their essential role during embryogenesis [65, 79]. Furthermore, mRpL55 was found to have a crucial role during the development of fruit flies, with mRpL55 null individuals showing increased mortality, reduced growth and activity after hatching [80]. Due to the nature of the bovine variant, two-thirds of the truncated protein are lacking, with most likely non-sense mediated decay taking place, thus it can be assumed that the variant leads to a loss-of-function of *MRPL55* in the homozygous state. Finally, this could explain putative early embryonic lethality and is consistent with the complete lack of homozygous animals. Moreover, the described biological effect is supported by the association study indicating negative associations with fertility traits and a reduction in heifer survival. The latter can be explained with fertility issues being one of the major culling reasons in dairy cattle [81, 82].

For the BS population, we propose three additional variants as candidate causal variants impairing postnatal survival. First, the BH24-associated missense variant in the *CPT1C* gene, which is known to be important for the proper development of the brain [83], was predicted to have a deleterious effect on protein level and effects on an evolutionary highly conserved nucleotide. Disorders associated with pathogenic *CPT1C* variants cause lethal spastic paraplegia (OMIM #616282). The BH24 haplotype shows significant associations with several birth traits and therefore this variant represents a plausible candidate for reduced homozygosity, which was confirmed after targeted genotyping. Second, the BH6-associated most likely pathogenic missense variant in

MARS2, a gene that plays an important role in embryonic development, is analogous to *MRPL55,* potentially leading to developmental arrest during early embryogenesis [79]. In humans, variants in this gene are associated with mitochondrial respiratory chain disorders and lead to retarded growth and hypotonia [84], and are also associated with a neurodevelopmental disorder (OMIM #611390) [68]. Third, the most likely pathogenic BH34-associated missense variant in *ACSL5*, which is involved in lipid metabolism, might cause a developmental delay due to disturbed fat metabolism in various mammalian species (OMIM #605677, OMIA #002226-9615) [67], and might explain the negative effects that we observed on growth-related traits. Both the *MARS2* and *ACSL5* variants need to be evaluated by further genotyping to confirm the postulated effects.

For the OB population, we propose two novel variants as candidate causal variants impairing embryonic survival. First, the OH4-associated frameshift variant in the *LIG3* gene, which is related to early embryonic lethality in mice [66], is of high interest. Functionally, LIG3 is a DNA binding ligase that is responsible for the repair of strand breaks [85, 86]. However, more precisely, the function of LIG3 is not essential for the repair of nuclear DNA since other proteins can compensate for the lack of LIG3, but it is essential for the repair of mitochondrial DNA [85, 86]. Second, the OH2-associated missense variant in the *TUBGCP5* gene, although not predicted as having a damaging effect on an evolutionary highly conserved nucleotide, it might represent a suitable candidate, since *TUBGCP5* plays an important role in the γ-tubulin ring complex which binds to the centrosome and thereby affects the cell cycle [87–89]. This complex has a highly conserved structure including six γ-tubulin proteins. TUBGCP5, in contrast to the other proteins, is present in single copy in the γ-tubulin ring complex [87–89]. In humans, an association between a *TUBGCP5* missense variant and microcephaly, which is a severe anomaly, has been shown [90]. Although to date these two variants have not been evaluated in the broader population, the observed significant negative effect of the OH2 haplotype on the maternal birth trait 'percentage of live births' and the deleterious nature of the *LIG3* and *TUBGCP5* variants support the assumed negative impact on reproduction.

Furthermore, three additional haplotypes (OH7, OH8, and OH9) which showed no significant HWE-deviation after correcting for FDR were detected. Interestingly, we found very convincing candidate variants for these haplotypes, such as the missense variant in the *HSD3B7* gene located on bovine chromosome 25. This variant is in high LD with OH9 and no homozygous carrier animals were found. *HSD3B7* is associated with the often lethal recessively inherited bile acid synthesis defect (BASD; OMIM #607764) and with progressive liver disease characterized by malabsorption of lipids and vitamins [91, 92]. *HSD3B7* knock-out mice showed increased mortality of homozygous newborns and decreased cholesterol absorption in the surviving animals [92]. Given the nature of the bovine *HSD3B7* missense variant and the observed negative effect on the growth-related trait 'bull survival', we speculate that homozygous mutant animals are most likely either non-viable or impaired in development.

Interestingly a unique haplotype region on bovine chromosome 11 was mapped in both populations, suggesting a common causative variant and therefore underlining the common heritage of the two Swiss Braunvieh cattle populations [34, 35]. Nevertheless, previous studies in OB have identified this region as a signature of selection [35, 93]. For these haplotypes (BH19 and OH3), we have not identified any potentially causative variant. The most probable variant identified was a missense variant in the *myomaker (MYMK)* gene; however, this variant occurs in the homozygous state in many animals. Therefore, it is likely a less suitable candidate for the depletion of homozygosity. Interestingly the OB-specific GWAS results for the growth-related carcass conformation score traits pinpoint the same region on chromosome 11 at 104 Mb. Based on the function of the *MYMK* gene on skeletal muscle development [94], we suggest that the impact of this variant on economically important growth traits needs to be subsequently evaluated. In contrast, deletion of *Mymk* in mice is perinatal lethal [94]. This example illustrates nicely the case of a variant that potentially underlies balancing selection as shown previously in pigs and cattle [95, 96]. Alternatively, and more likely, other undetected variants, such as more complex structural variants, might be responsible for the observed depletion in homozygosity in this genomic region.

As described above, not all the variants that we claim to be potentially causing disease could be evaluated in the current population, as many of them were unfortunately not included on the custom array. There are technical reasons why many variants were not included on this final array, such as interference with adjacent variants, high GC-content, repetitive DNA, multi-allelic variation, etc. Nonetheless, the SWISScow custom chip proved to be a very powerful and efficient method to validate the segregation of the most likely causative variants in the current population. Furthermore, the genomic data of a population changes over time since the haplotype frequencies can change due to the impact of popular sires. Regarding the haplotype regions, we may have failed to detect some specific haplotype regions because the window size was set at 50 markers, since Hoff et al.

[25] showed that the homozygosity rate negatively correlates with window length.

Although we propose 24 new variants, including six of high importance, the causal variants for numerous haplotype regions including BH1 remain unclear. For example, in the BS population, we identified the highly HWE-deviating haplotype BH23 on chromosome 17, which showed significant haplotype associations with growth-related and birth traits, plus suggestive associations with fertility. In parallel, for the same region, suggestive GWAS associations with the fertility trait 'non-return rate' were observed, but we failed to identify candidate causal variants.

This analysis is based on the latest reference sequence ARS-UCD1.2 of a Hereford cow, however, the reference sequence still spans many gaps and includes more than 2000 unplaced scaffolds that could potentially harbor important coding sequences [37]. Recently, it was shown that reference graphs based on sequence data of several breeds can include new functional sequences [97]. Another drawback of our approach is that we restricted our analysis to coding variants, while non-coding variants impairing the expression of a protein or more complex structural variants disturbing the function of genes might also be causative. The interference of non-detected structural variants could also be an explanation for the rather low LD values in some regions [98, 99]. Furthermore, as a few homozygous diplotype carriers are alive, some haplotypes may arise due to the high selective pressure on traits within the breeding program. As shown by the example of cholesterol deficiency in Holstein cattle, the associated haplotype can segregate in the population under two indistinguishable versions, a derived variant and its ancestral version [100]. In summary, the reported haplotype regions might reflect population-specific characteristics of the breeding programs, such as the intensive use of individual bulls by artificial insemination or by natural service, a common practice in Swiss OB. Most likely, the focus applied on monogenic Mendelian disorders is too simplistic since a polygenic architecture can often be assumed. The latter is supported by the GWAS results that highlight numerous additional associated genomic regions, which are not affected by a depletion of homozygosity.

Conclusions

For the first time, we applied the trio-based mapping approach in cattle for the genome-wide detection of haplotypes showing reduced homozygosity that segregate at low to moderate frequencies. We present a short list of potentially causative variants and highlight four coding variants for the Brown Swiss population and two for the Original Braunvieh population located in candidate genes that are involved in embryonic and pre-weaning lethality. Although it will be challenging to evaluate all the candidate variants that we propose here by targeted monitoring of at risk matings and possible clinical examination of live homozygotes, it is important to rule out causality. This study illustrates the difficulty to select for improved fertility and rearing success by focusing on monogenic disorders, since our GWAS results confirmed the polygenic nature of these traits. Nonetheless, the proposed candidate causative variants will help to refine DNA-based selection decisions to improve female fertility and rearing success in Swiss Braunvieh cattle.

Abbreviations

ACSL5: *Acyl-CoA synthetase long chain family member 5*; Arg: Arginine; BASD: bile acid synthesis defect; BH: Brown Swiss Haplotype; BH1: Brown Swiss Haplotype 1; BH2: Brown Swiss Haplotype 2; bp: base pairs; BS: Brown Swiss; *CPT1C*: *Carnitine palmitoyltransferase 1C*; Cys: Cysteine; drEBV: deregressed estimated breeding value; FDR: False discover rate; FH2: Fleckvieh Haplotype 2; GS: genomic selection; GWAS: genome-wide association study; HD: high density SNP panel; HO: Holstein; *HSD3B7*: *Hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7*; HWE: Hardy–Weinberg equilibrium; LD: low density SNP panel; *LIG3*: *DNA ligase 3*; *MARS2*: *Methionyl-tRNA synthetase 2, mitochondrial*; Mb: Mega bases; *MRPL55*: *Mitochondrial ribosomal protein L55*; *MYMK*: *Myomaker*; OB: dual-purpose breed Original Braunvieh; OH: Original Braunvieh Haplotype; pgp: parent – grant parent (paternal half-sib groups); QTL: quantitative trait loci; r2: R-squared value of the LD analysis; *SCFD2*: *Sec1 family domain containing 2*; *SLC2A2*: *Solute carrier family 2 member 2*; SNV: small nucleotide variants; SNP: single nucleotide polymorphism; SNP array: genotyping arrays including single nucleotide polymorphisms; *TCF3*: *Transcription factor 3*; TRD: transmission ratio distortion; *TUBD1*: *Tubulin delta 1*; *TUBGCP5*: *Tubulin gamma complex associated protein 5*; VCF: variant call format file; WGS: whole-genome sequencing.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12711-021-00686-3.

Additional file 1: Table S1. WGS sample information. All whole-genome sequencing samples included in this study and their project and sample ids from the European Nucleotide Archive (ENA). In addition, the sex, breed and average read depth and insert size are provided.

Additional file 2: Table S2. Haplotype information. All haplotypes for the trio- and pgp-based approaches for the BS (**A**) and the OB population (**B**). The descriptive data include the type of analysis, genomic positions, number of expected and observed animals, haplotype frequency, p-value from exact tests of Hardy–Weinberg equilibrium and the adjusted p-value according to the Benjamini-Yekutieli procedure.

Additional file 3: Table S3. Output of the haplotype analyses. Results of the estimation of the haplotype effect with different conformation, production and reproduction traits for the BS (**A**) and OB populations (**B**).

Additional file 4: Figure S1. Manhattan plots and their QQ-plots of the GWAS results for the BS and OB populations. There is a page for each of the fertility, birth and growth-related trait groups, including a Manhattan plot for every single trait according to Table 2.

Additional file 5: Table S4. GWAS results. Significant ($p < 4.35e{-}7$) GWAS results for the BS (**A**) and (**B**) OB populations.

Additional file 6: Table S5. Comprehensive list of candidate causal variants. The data for BS (**A**) and OB (**B**) are provided, with information regarding haplotypes and the potential candidate causal variants based on the

reference sequence ARS-UCD1.2 [36] and the NCBI Annotation Release 106 [54], including variant effect predictions (siftScore [63] and MutPred2 score [64]) and base conservation scores (phyloP and phastCons [61, 62]). Furthermore, the genotype distributions across whole-genome sequencing and genotyping data are included.

Acknowledgements
The authors are grateful to the Swiss cattle breeding association Braunvieh Schweiz for providing the genomic SNP array and the phenotypic data. We are grateful to Swissgenetics and SelectStar for sending sperm samples of the sires. The assistance of Nathalie Besuchet-Schmutz with DNA extraction is acknowledged. We acknowledge the assistance of the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome re-sequencing experiments. Furthermore, we thank the Interfaculty Bioinformatics Unit of the University of Bern for providing high performance computing infrastructure. The 1000 Bull Genomes Project consortium (Christy Vander Jagt, Hans Daetwyler) is acknowledged for providing the variant catalogue. Christine F. Baes assisted in proofreading.

Authors' contributions
IMH prepared and analyzed the whole-genome sequencing data, analyzed and interpreted all the obtained results. Furthermore, IMH drafted this manuscript including tables and visualizations. CD acquired the whole-genome sequencing data. MS and FRS handled and analyzed the SNP genotyping data, including the genome-wide association studies and haplotype analysis. CD and FRS designed and supervised the project. All authors read and approved the final manuscript.

Funding
This study was funded by the Swiss National Science Foundation, grant number 172911. Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, 172911, Cord Drögemüller

Availability of data and materials
For access to the genotyping data, interested people are asked to contact the Braunvieh breeding association directly.

Declarations

Ethics approval and consent to participate
According to the Swiss laws, no ethics approval was needed. In this study, already available data was reused and mined by using novel approaches.

Consent for publication
Not applicable.

Competing interests
The authors declare that there have no competing interests.

Author details
[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland. [2]Qualitas AG, 6300 Zug, Switzerland.

Received: 11 May 2021 Accepted: 26 November 2021
Published online: 16 December 2021

References

1. Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF. A 100-year review: Identification and genetic selection of economically important traits in dairy cattle. J Dairy Sci. 2017;100:10251–71.
2. Pryce JE, Woolaston R, Berry DP, Wall E, Winters M, Butler R, et al. World trends in dairy cow fertility. In Proceedings of the 10th World Congress on Genetics Applied to Livestock Production. 18–22 August 2014; Vancouver. 2014.
3. Walsh SW, Williams EJ, Evans ACO. A review of the causes of poor fertility in high milk producing dairy cows. Anim Reprod Sci. 2011;123:127–38.
4. Pryce JE, Royal MD, Garnsworthy PC, Mao IL. Fertility in the high-producing dairy cow. Livest Prod Sci. 2004;86:125–35.
5. Goddard ME, Hayes BJ, Meuwissen THE. Genomic selection in livestock populations. Genet Res (Camb). 2010;92:413–21.
6. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci. 2009;92:433–43.
7. Casellas J, Cañas-Alvarez JJ, González-Rodríguez A, Puig-Oliveras A, Fina M, Piedrafita J, et al. Bayesian analysis of parent-specific transmission ratio distortion in seven Spanish beef cattle breeds. Anim Genet. 2017;48:93–6.
8. Casellas J, Gularte RJ, Farber CR, Varona L, Mehrabian M, Schadt EE, et al. Genome scans for transmission ratio distortion regions in mice. Genetics. 2012;191:247–59.
9. Pausch H, Schwarzenbacher H, Burgstaller J, Flisikowski K, Wurmser C, Jansen S, et al. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. BMC Genomics. 2015;16:312.
10. Cole JB, VanRaden PM, Null DJ, Hutchison JL, Cooper TA, Hubbard SM. Haplotype tests for recessive disorders that affect fertility and other traits. Animal Improvement Program Research Report. 2018. https://www.aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html. Accessed 2021 Apr 15.
11. Schwarzenbacher H, Burgstaller J, Seefried FR, Wurmser C, Hilbe M, Jung S, et al. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC Genomics. 2016;17:400.
12. Sonstegard TS, Cole JB, VanRaden PM, Van Tassell CP, Null DJ, Schroeder SG, et al. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. PLoS One. 2013;8:e54872.
13. Charlier C, Agerholm JS, Coppieters W, Karlskov-Mortensen P, Li W, de Jong G, et al. A deletion in the bovine FANCI gene compromiss fertility by causing fetal death and brachyspina. PLoS One. 2012;7:e43085.
14. McClure M, Kim E, Bickhart D, Null D, Cooper T, Cole J, et al. Fine mapping for weaver syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. PLoS One. 2013;8:e59251.
15. Schütz E, Wehrhahn C, Wanjek M, Bortfeld R, Wemheuer WE, Beck J, et al. The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. PLoS One. 2016;11:e0154602.
16. Null DJ, Huthcins JL, Bickhart DM, VanRaden PM, Cole JB. Discovery of a haplotype affecting fertility in Ayrshire dairy cattle and identification of a putative causal variant. J Dairy Sci. 2017;100:S199.
17. Guarini AR, Sargolzaei M, Brito LF, Kroezen V, Lourenco DAL, Baes CF, et al. Estimating the effect of the deleterious recessive haplotypes AH1 and AH2 on reproduction performance of Ayrshire cattle. J Dairy Sci. 2019;102:5315–22.
18. Cooper TA, Wiggans GR, Null DJ, Hutchison JL, Cole JB. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. J Dairy Sci. 2014;97:3878–82.
19. VanRaden PM, Olson KM, Null DJ, Hutchison JL. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J Dairy Sci. 2011;94:6153–61.
20. Ben Braiek M, Fabre S, Hozé C, Astruc J-M, Moreno-Romieux C. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. Genet Sel Evol. 2021;53:41.
21. Derks MFL, Megens HJ, Bosse M, Lopes MS, Harlizius B, Groenen MAM. A systematic survey to identify lethal recessive variation in highly managed pig populations. BMC Genomics. 2017;18:858.
22. Derks MFL, Megens HJ, Bosse M, Visscher J, Peeters K, Bink M, et al. A survey of functional genomic variation in domesticated chickens. Genet Sel Evol. 2018;50:17.
23. Georges M, Charlier C, Hayes B. Harnessing genomic information for livestock improvement. Nat Rev Genet. 2019;20:135–56.

24. Abdalla EA, Id-Lahoucine S, Cánovas A, Casellas J, Schenkel FS, Wood BJ, et al. Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. Anim Genet. 2020;51:876–89.
25. Hoff JL, Decker JE, Schnabel RD, Taylor JF. Candidate lethal haplotypes and causal mutations in Angus cattle. BMC Genomics. 2017;18:799.
26. Mesbah-Uddin M, Hoze C, Michot P, Barbat A, Lefebvre R, Boussaha M, et al. A missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. J Dairy Sci. 2019;102:6340–56.
27. Fritz S, Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, et al. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. PLoS One. 2013;8:e65550.
28. Michot P, Fritz S, Barbat A, Boussaha M, Deloche MC, Grohs C, et al. A missense mutation in PFAS (phosphoribosylformylglycinamidine synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbeliarde dairy cattle. J Dairy Sci. 2017;100:8176–87.
29. Fritz S, Hoze C, Rebours E, Barbat A, Bizard M, Chamberlain A, et al. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. J Dairy Sci. 2018;101:6220–31.
30. Hozé C, Escouflaire C, Mesbah-Uddin M, Barbat A, Boussaha M, Deloche MC, et al. Short communication: A splice site mutation in CENPU is associated with recessive embryonic lethality in Holstein cattle. J Dairy Sci. 2020;103:607–12.
31. Kunz E, Rothammer S, Pausch H, Schwarzenbacher H, Seefried FR, Matiasek K, et al. Confirmation of a non-synonymous SNP in PNPLA8 as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. Genet Sel Evol. 2016;48:21.
32. Charlier C, Li W, Harland C, Littlejohn M, Coppieters W, Creagh F, et al. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. Genome Res. 2016;26:1333–41.
33. McClure MC, Bickhart D, Null D, VanRaden P, Xu LY, Wiggans G, et al. Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative causative mutation in SMC2 for HH3. PLoS One. 2014;9:e92769.
34. Hagger C. Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. J Anim Breed Genet. 2005;122:405–13.
35. Signer-Hasler H, Burren A, Neuditschko M, Frischknecht M, Garrick D, Stricker C, et al. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. Genet Sel Evol. 2017;49:83.
36. National Center for Biotechnology Information. ARS-UCD1.2 Assembly. https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1. Accessed 2021 Apr 15.
37. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9:giaa021. https://doi.org/10.1093/gigascience/giaa021.
38. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. BMC Genomics. 2014;15:478.
39. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81:1084–97.
40. Sargolzaei M. SNP1101 User's Guide. Version 1.0. Guelph: HiggsGene Solutions Inc.; 2014.
41. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet. 2005;76:887–93.
42. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 2001;29:1165–88.
43. Garrick DJ, Taylor JF, Fernando RL. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genet Sel Evol. 2009;41:55.
44. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76–82.
45. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.
46. Turner S. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J Open Source Softw. 2018;3:1–2.
47. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
48. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
49. Broad Institute. Picard Tools version 2.18.2. http://broadinstitute.github.io/picard/. 2018. Accessed 2021 Apr 15.
50. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
51. Bovine Research Community. 1000 Bull Genomes Project. http://www.1000bullgenomes.com/. Accessed 2018 Jun 15.
52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
53. Brentp/goleft. Github. https://github.com/brentp/goleft. Accessed 2018 Aug 20.
54. National Center for Biotechnology Information. NCBI Bos taurus Annotation Release 106. https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/. Accessed 2018 Aug 20.
55. Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, et al. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program. SnpSift Front Genet. 2012;3:35.
56. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. Annu Rev Anim Biosci. 2019;7:89–102.
57. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
58. Hu Y, Yan C, Hsu CH, Chen QR, Niu K, Komatsoulis GA, et al. OmicCircos: a simple-to-use R package for the circular visualization of multidimensional Omics data. Cancer Inform. 2014;13:13–20.
59. Kuhn RM, Haussler D, Kent JW. The UCSC genome browser and associated tools. Brief Bioinform. 2013;14:144–61.
60. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921.
61. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of non-neutral substitution rates on mammalian phylogenies. Genome Res. 2010;20:110–21.
62. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005;15:1034–50.
63. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and Indels. PLoS One. 2012;7:e46688.
64. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam HJ, et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat Commun. 2020;11:5918.
65. Cheong A, Lingutla R, Mager J. Expression analysis of mammalian mitochondrial ribosomal protein genes. Gene Expr Patterns. 2020;38:119147.
66. Puebla-Osorio N, Lacey DB, Alt FW, Zhu C. Early embryonic lethality due to targeted inactivation of DNA ligase III. Mol Cell Biol. 2006;26:3935–41.
67. Bowman TA, O'Keeffe KR, D'Aquila T, Yan QW, Griffin JD, Killion EA, et al. Acyl CoA synthetase 5 (ACSL5) ablation in mice increases energy expenditure and insulin sensitivity and delays fat absorption. Mol Metab. 2016;5:210–20.
68. Bayat V, Thiffault I, Jaiswal M, Tétreault M, Donti T, Sasarman F, et al. Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol. 2012;10:e1001288.
69. El-Hamidi M, Leipold HW, Vestweber JGE, Saperstein G. Spinal muscular atrophy in Brown Swiss calves. Zentralbl Veterinarmed A. 1989;36:731–8.
70. Krebs S, Medugorac I, Röther S, Strässer K, Förster M. A missense mutation in the 3-ketodihydrosphingosine reductase FVT1 as candidate causal mutation for bovine spinal muscular atrophy. Proc Natl Acad Sci USA. 2007;104:6746–51.
71. Leipold HW, Blaugh B, Huston K, Edgerly CG, Hibbs CM. Weaver syndrome in Brown Swiss cattle: clinical signs and pathology. Vet Med Small Anim Clin. 1973;68:645–7.

72. Joller S, Stettler M, Locher I, Dettwiler M, Seefried F, Meylan M, et al. Fanconi-bickel-syndrom: A novel genetic disease in Original Braunvieh. Schweiz Arch Tierheilkd. 2018;160:179–84.
73. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B cell formation. Cell. 1994;79:875–84.
74. Liu W, Liu Y, Guo T, Hu C, Luo H, Zhang L, et al. TCF3, a novel positive regulator of osteogenesis, plays a crucial role in miR-17 modulating the diverse effect of canonical Wnt signaling in different microenvironments. Cell Death Dis. 2013;4:e539.
75. Narasimhan VM, Xue Y, Tyler-Smith C. Human knockout carriers: dead, diseased, healthy, or improved? Trends Mol Med. 2016;22:341–51.
76. Nosková A, Bhati M, Kadri NK, Crysnanto D, Neuenschwander S, Hofer A, et al. Characterization of a haplotype-reference panel for genotyping by low-pass sequencing in Swiss Large White pigs. BMC Genomics. 2021;22:290.
77. Ros-Freixedes R, Gonen S, Gorjanc G, Hickey JM. A method for allocating low-coverage sequencing resources by targeting haplotypes rather than individuals. Genet Sel Evol. 2017;49:78.
78. Mai N, Chrzanowska-Lightowlers ZMA, Lightowlers RN. The process of mammalian mitochondrial protein synthesis. Cell Tissue Res. 2017;367:5–20.
79. Cheong A, Archambault D, Degani R, Iverson E, Tremblay KD, Mager J. Nuclear-encoded mitochondrial ribosomal proteins are required to initiate gastrulation. Development. 2020. https://doi.org/10.1242/dev.188714.
80. Tselykh TV, Roos C, Heino TI. The mitochondrial ribosome-specific MrpL55 protein is essential in Drosophila and dynamically required during development. Exp Cell Res. 2005;307:354–66.
81. Seegers H, Beaudeau F, Fourichon C, Bareille N. Reasons for culling in French Holstein cows. Prev Vet Med. 1998;36:257–71.
82. Rilanto T, Reimus K, Orro T, Emanuelson U, Viltrop A, Mõtus K. Culling reasons and risk factors in Estonian dairy cows. BMC Vet Res. 2020;16:173.
83. Carrasco P, Jacas J, Sahún I, Muley H, Ramírez S, Puisac B, et al. Carnitine palmitoyltransferase 1C deficiency causes motor impairment and hypoactivity. Behav Brain Res. 2013;256:291–7.
84. Webb BD, Wheeler PG, Hagen JJ, Cohen N, Linderman MD, Diaz GA, et al. Novel, compound heterozygous, single-nucleotide variants in MARS2 associated with developmental delay, poor growth, and sensorineural hearing loss. Hum Mutat. 2015;36:587–92.
85. Simsek D, Furda A, Gao Y, Artus J, Brunet E, Hadjantonakis A-K, et al. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. Nature. 2011;471:245–8.
86. Gao Y, Katyal S, Lee Y, Zhao J, Rehg JE, Russell HR, et al. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. Nature. 2011;471:240–4.
87. Murphy SM, Preble AM, Patel UK, O'Connell KL, Dias DP, Moritz M, et al. GCP5 and GCP6: two new members of the human gamma-tubulin complex. Mol Biol Cell. 2001;12:3340–52.
88. Izumi N, Fumoto K, Izumi S, Kikuchi A. GSK-3β regulates proper mitotic spindle formation in cooperation with a component of the γ-tubulin ring complex, GCP5. J Biol Chem. 2008;283:12981–91.
89. Farache D, Emorine L, Haren L, Merdes A. Assembly and regulation of γ-tubulin complexes. Open Biol. 2018;8:170266.
90. Maver A, Čuturilo G, Kovanda A, Miletić A, Peterlin B. Rare missense TUBGCP5 gene variant in a patient with primary microcephaly. Eur J Med Genet. 2019;62:103598.
91. Cheng JB, Jacquemin E, Gerhardt M, Nazer H, Cresteil D, Heubi JE, et al. Molecular genetics of 3β-hydroxy-Δ5-C27-steroid oxidoreductase deficiency in 16 patients with loss of bile acid synthesis and liver disease. J Clin Endocrinol Metab. 2003;88:1833–41.
92. Shea HC, Head DD, Setchell KDR, Russell DW. Analysis of HSD3B7 knockout mice reveals that a 3α-hydroxyl stereochemistry is required for bile acid function. Proc Natl Acad Sci USA. 2007;104:11526–33.
93. Bhati M, Kadri NK, Crysnanto D, Pausch H. Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. BMC Genomics. 2020;21:27.
94. Millay DP, O'Rourke JR, Sutherland LB, Bezprozvannaya S, Shelton JM, Bassel-Duby R, et al. Myomaker is a membrane activator of myoblast fusion and muscle formation. Nature. 2013;499:301–5.
95. Derks MFL, Lopes MS, Bosse M, Madsen O, Dibbits B, Harlizius B, et al. Balancing selection on a recessive lethal deletion with pleiotropic effects on two neighboring genes in the porcine genome. PLoS Genet. 2018;14:e1007661.
96. Kadri NK, Sahana G, Charlier C, Iso-Touru T, Guldbrandtsen B, Karim L, et al. A 660-Kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic Red cattle: Additional evidence for the common occurrence of balancing selection in livestock. PLoS Genet. 2014;10:e1004049.
97. Crysnanto D, Leonard AS, Fang Z-H, Pausch H. Novel functional sequences uncovered through a bovine multi-assembly graph. Proc Natl Acad Sci USA. 2021;118:e2101056118.
98. Posey JE. Genome sequencing and implications for rare disorders. Orphanet J Rare Dis. 2019;14:153.
99. Bocher O, Génin E. Rare variant association testing in the non-coding genome. Hum Genet. 2020;139:1345–62.
100. Menzi F, Besuchet-Schmutz N, Fragnière M, Hofstetter S, Jagannathan V, Mock T, et al. A transposable element insertion in *APOB* causes cholesterol deficiency in Holstein cattle. Anim Genet. 2016;47:253–7.
101. Häfliger IM, Marchionatti E, Stengård M, Wolf-Hofstetter S, Paris JM, Jacinto JG, et al. *CNGB3* missense variant causes recessive achromatopsia in Original Braunvieh cattle. Int J Mol Sci. 2021;22:12440.

Publisher's Note

Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle

Journal:	Scientific Reports
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, visualizations, writing original draft and revised versions
Displayed version:	published version
DOI:	10.1038/s41598-022-09403-6

scientific reports

Check for updates

OPEN

Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle

Irene M. Häfliger[1]✉, Mirjam Spengeler[2], Franz R. Seefried[2] & Cord Drögemüller[1]

Mendelian variants can determine both insemination success and neonatal survival and thus influence fertility and rearing success of cattle. We present 24 deficient homozygous haplotype regions in the Holstein population of Switzerland and provide an overview of the previously identified haplotypes in the global Holstein breed. This study encompasses massive genotyping, whole-genome sequencing (WGS) and phenotype association analyses. We performed haplotype screenings on almost 53 thousand genotyped animals including 114 k SNP data with two different approaches. We revealed significant haplotype associations to several survival, birth and fertility traits. Within haplotype regions, we mined WGS data of hundreds of bovine genomes for candidate causal variants, which were subsequently evaluated by using a custom genotyping array in several thousand breeding animals. With this approach, we confirmed the known deleterious *SMC2*:p.Phe1135Ser missense variant associated with Holstein haplotype (HH) 3. For two previously reported deficient homozygous haplotypes that show negative associations to female fertility traits, we propose candidate causative loss-of-function variants: the HH13-related *KIR2DS1:*p.Gln159* nonsense variant and the HH21-related *NOTCH3:*p.Cys44del deletion. In addition, we propose the *RIOX1:*p.Ala133_Glu142del deletion as well as the *PCDH15:*p.Leu867Val missense variant to explain the unexpected low number of homozygous haplotype carriers for HH25 and HH35, respectively. In conclusion, we demonstrate that with mining massive SNP data in combination with WGS data, we can map several haplotype regions and unravel novel recessive protein-changing variants segregating at frequencies of 1 to 5%. Our findings both confirm previously identified loci and expand the spectrum of undesired alleles impairing reproduction success in Holstein cattle, the world's most important dairy breed.

Holstein is by far the most popular breed of cattle in the world, bred and raised for its high milk yield[1]. Especially with the improved reproductive technologies that arose in the last century, the genomic exchange over the world increased and led to the use of a limited number of prominent sires across the globe[1–3]. For Holstein cattle the difference between the absolute population size and the effective population size differs highly in all parts of the world and the inbreeding is increasing after the recent implementation of genomic selection[4]. With the effective population size being much smaller, the decrease in genetic variation and the increase in inbreeding is accelerated[2]. There has been a recognisable decline in reproductive performance in high producing dairy cattle[5,6], which could be shown to be due to artificial selection on production traits and negative hitchhiking effects[3]. Since fertility problems are one of the most common reasons for culling cattle[7,8] and animals of the Holstein breed are more likely to be culled than animals of other breeds[8], there is a high need to improve female fertility, especially in Holstein cattle.

Genetic analyses of female fertility range from genome-wide association studies (GWAS) identifying genomic regions associated with fertility traits[9–12], to the identification of genomic regions showing reduced homozygosity due to recessive mostly embryonic lethal variants[13] and to the investigation of effects of specific variants in functional candidate genes[14,15]. Due to the routinely used single nucleotide polymorphism (SNP) array genotyping implemented in breeding programs about a decade ago, comprehensive population-wide genomic data of the current breeding populations are available. VanRaden et al. were the first to propose to screen this kind of data to identify genomic regions with fewer homozygous animals than expect. In their article they listed five haplotypes deviating significantly form the Hardy–Weinberg equilibrium (HWE) and thereby potentially harbouring lethal variants in the American Holstein population (Table S1)[13]. Three of these haplotypes, namely Holstein haplotypes 1, 2 and 3 (HH1-HH3), showed significant negative associations to reproduction traits. Interestingly, two

[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland. [2]Qualitas AG, 6300 Zug, Switzerland. ✉email: irene.haefliger@vetsuisse.unibe.ch

Haplotype name	Associated disorder	Gene	OMIA	Variant description						Publications
				chr	Position[a]	Description	Transcript of interest[b]	Coding DNA change	Protein change	
HH0	Brachyspina	*FANCI*	000151-9913	21	21,184,869–21,188,198	Gross deletion (frameshift)		3.3 kb deletion		19,20,27,85,86
HH1	Embryonic lethality	*APAF1*	000001-9913	5	62,810,245	SNV (nonsense)	NM_001191507.1	c.1702C > T	p.Gln568*	13,16,22,85
HH2	Embryonic lethality	*IFT80*	001823-9913	1	107,172,616	SNV (frameshift)	XM_024984168.1	c.747delT	p.Leu250fs	13,24,28,85
HH3	Abortion	*SMC2*	001,824-9913	8	93,753,358	SNV (missense)	XM_015472668.2	c.3404T > C	p.Phe1135Ser	13,18,19,85,87
HH4	Abortion	*GART*	001826–9913	1	1,997,582	SNV (missense)	NM_001040473.2	c.869A > C	p.Asn290Thr	16,27,85
HH5	Abortion	*TFB1M*	001941-9913	9	93,223,651–93,370,998	Gross deletion		139 kb deletion		26,27,85
HH6	Embryonic lethality	*SDE2*	002149-9913	16	29,020,700	SNV (start-lost)	NM_001099065.2	c.2T > C	p.Met1?	27
HH7	Embryonic lethality	*CENPU*	001830-9913	27	15,123,636	5 bp deletion (splice site)	XM_002698654.5	c.15123637_15123640delTTACT		17
HHB	Bovine leukocyte adhesion deficiency (BLAD)	*ITGB2*	000595-9913	1	144,770,078	SNV (missense)	NM_175781.1	c.383A > G	p.Asp128Gly	51,85
HHC	Complex vertebral malformation (CVM)	*SLC35A3*	001340-9913	3	43,261,945	SNV (missense)	NM_001105386.1	c.538G > T	p.Val180Phe	21,85,88
HHD	Deficiency of uridine monophosphate synthase (DUMPS)	*UMPS*	000262-9913	1	69,151,931	SNV (nonsense)	NM_177508.1	c.1213C > T	p.Arg405*	85,89,90
CDH	Cholesterol deficiency; juvenile mortality	*APOB*	001965-9913	11	77,891,733	Large insertion (frameshift)		ERV insertion		26,29,30,53,85

Table 1. Previously identified lethal haplotype regions and their associated candidate causal variants in Holstein cattle. [a]According to the reference sequence ARS-UCD1.2[31]. [b]According to the NCBI Annotation Release 106[91].

subsequently performed similar analyses in the French Holstein population revealed 14 additional haplotype regions while confirming the three previously described haplotypes HH1, 2 and 3 (Table S1)[16,17]. Therefore, performing new analyses regularly as genotyping data are accumulating was recommended[17]. In two similar screenings of the Nordic Holstein population another nineteen haplotype regions in addition to the previously reported HH3 region and the *FANCI*-related brachyspina-associated haplotype HH0, originally designated as HBY, were identified by Sahana et al. and Wu et al. (Table S1)[18,19]. In the Danish Holstein population a haplotype associated with a large genomic deletion representing a loss-of-function of the *FANCI* gene causing brachyspina[20], as well as a haplotype associated with a missense variant in the *SLC35A3* gene causing complex vertebral malformation[21] were identified (Table 1, Table S1). The latter two variants were initially thought to cause congenital malformation, but it was shown that in homozygous state, both mutations typically cause death in utero. Therefore, the two associated haplotypes (HH0 and HHC) were included in the list of 43 known Holstein haplotypes with reduced homozygosity (Table S1).

So far, candidate causal variants for seven of these Holstein haplotypes were identified either by whole-genome or whole-exome sequencing affecting different genes of importance for reproduction and/or development: *APAF1* (HH1), *IFT80* (HH2), *SMC2* (HH3), *GART* (HH4), *TFB1M* (HH5), *SDE2* (HH6) and *CENPU* (HH7) (Table 1)[16,17,22–28]. In addition, a haplotype associated with increased juvenile mortality due to cholesterol deficiency (CDH) in the German Holstein population[29] had been detected to be associated with a transposable element insertion in in the *APOB* gene that causes the metabolic disorder (Table 1, Table S1)[26,29,30].

Taken together, these findings indicate, that even though Holstein is an international breed characterized by a limited number of founding animals, the national subpopulations carry a different genetic load due to local selection decisions. Therefore, the aim of this study was to explore the accumulated massive genomic and phenotypic data of the local Swiss Holstein cattle population to identify genomic regions impairing fertility and rearing success. We applied a combination of methods including screening for missing homozygosity based on SNP-genotyped trios, trait association studies using estimated breeding values for female fertility traits and linkage disequilibrium analyses exploiting both SNP and whole-genome sequencing (WGS) data. The ultimate

Trait group	Trait sub-group	Trait	Description	Abbreviation
Fertility traits	Fertility	Non-return rate heifer	Heifers non-return rate after 56 days, binary	NRh
		Non-return rate cow	Cows non-return rate after 56 days, binary	NRc
		Interval first to last insemination heifer	Interval between first and last insemination for heifer, days	IFLh
		Interval first to last insemination cow	Interval between first and last insemination for cows, days	IFLc
		Interval calving to insemination	Interval from calving to first service, days	DFS
Birth traits	Birth history direct	Percentage normal births	Calving ease, scored between 1-without help to 5-dystocia	CEd
		Percentage live births	Percentage of calves born alive	SBd
		Birth weight	Weight of calve at birth, kg	BWd
		Gestation length	Days from successfull insemination to birth	GLd
		Multiple birth	Percentage of multiple births	MBd
	Birth history maternal	Percentage normal births	Calving ease, scored between 1-without help to 5-dystocia	CEm
		Percentage live births	Percentage of calves born alive	SBm
		Birth weight	Weight of calve at birth, kg	BWm
		Gestation length	Days from successfull insemination to birth	GLm
		Multiple birth	Percentage of multiple births	MBm
Survival traits	Rearing success	Survival period 1	Survival from day 3 up to 30th day of life	P1
		Survival heifer period 2	Survival of heifers from day 31 up to 458 days	P2h
		Survival bull period 2	Survival of young bulls from 31 up to 183 days	P2b

Table 2. Traits of interest.

goal was to pinpoint potentially causative protein-coding variants explaining reproductive failure due to recessive Mendelian disorders.

Materials and methods

The SNP array data of the Holstein population used was provided by the breeding association's swissherdbook and Holstein Switzerland. The quality-controlled (MAF > 0.01 and call rates > 0.9 per SNP and > 0.8 per animal) SNP data included 114 890 SNP combining a variety of SNP arrays with densities ranging from 3 to 150 k. SNP positions have been updated to the latest cattle reference sequence ARS-UCD1.2[31,32]. The software Fimpute v2.2[33] was used to impute the data in order to correct for false genotypes as well as to increase the SNP data of lower density genotyped animals. The comprehensive data included 52,961 Holstein cattle. From these data, two analysis data sets were formed: the trio-based analysis, where the complete trio was genotyped (sire, dam and offspring), with 17,915 animals and the pgp-based analysis, where the dam is replaced by the maternal grandsire (sire, maternal grandsire and offspring), with 30,315 animals. While the trio-based dataset allows tracking the direct inheritance of alleles, the pgp-based dataset includes some uncertainty. Nevertheless, the pgp-based approach, where all male animals are genotyped, represents more accurately the common genotyping scheme of current breeding programs, as can be seen from the higher number of animals in the dataset. Both datasets include animals born in Switzerland after 2009 and their ancestors' genotypes.

The described data sets (trio and pgp) were used in the software snp1101[34] to identify haplotypes with a significant deviation from HWE, including a correction of false discovery rate according to Benjamini-Yekutieli[35]. The chosen window size for the sliding-window approach was 50 SNP, based on the work of Hoff et al.[36]. The snp1101 output provides a list with all significant haplotypes, the number of observed and expected homozygous carriers, as well as the allele frequency and the haplotype itself. For all regions of interest, where the observed number of homozygous animals were lower than the number of expected homozygous animals, we have chosen the most significant haplotype (lowest p-value) to reduce the number of false positive carriers. Further, we predicted the diplotypes for these haplotypes in the entire genotyped Holstein population and selected carrier animals for WGS.

Association studies were conducted using the predicted diplotypes and all female reproduction traits from the Swiss routine genetic evaluation system. In total 18 traits were analysed, while traits can be grouped into fertility, birth and survival traits (Table 2). The estimated breeding values (EBV) were deregressed and used as pseudo-phenotypes in linear mixed models (LMM) in the GCTA software[37,38]. The genomic relationship matrix estimated using the method GRM of the GCTA software was included in the model to account for population structure. In addition, for the SNP data described above genome-wide association studies (GWAS) together with

deregressed EBVs as pseudo-phenotypes were performed. These single SNP regressions were implemented using the software snp1101[34], the genomic relationship matrix was calculated with the method proposed by VanRaden[39] and a Bonferroni-correction was applied ($p < 0.0000044$).

Whole-genome sequencing (WGS) was performed in an attempt to sequence animals potentially carrying variants causing reduced homozygosity in the identified genomic regions. Therefore, 37 Holstein animals were selected based on their diplotype status for WGS. Additionally, 656 genomes from other projects were available, leading to 691 publicly available genomes including 244 purebred Holstein cattle (Table S2). Sequence data preparations was done according to previously described methods[40], with the only difference in the recalibration step, where the known variants from the 1000 Bull Genomes Project run 7 (BQSR file version 2) were used[41]. This resulted in a variant call format (VCF) file encompassing all single-nucleotide variants (SNV) and short indels from the 691 animals. The average read depth per genome were calculated using covstat from goleft v0.1.19[42] and varied from 3.3× to 72.8×, with an average of 18.7× read depth (Table S1).

In order to design a custom SNP array, we selected candidate SNVs and short indels within the genomic regions of interest defined by the determined haplotype regions including plus/minus 2 Mb flanking sequence up- and downstream. Due to the variation in read depth, missing values and a single homozygous carrier animal were allowed, GATK recommended quality measures for hard filtering needed to be passed and at least a single Holstein animal needed to be carrier of a variant, but no more than 75% of all animals for it to be selected for the array design. Variants were included in a custom array using Axiom Microarray Genotyping Technology, designed under the umbrella of Swiss routine genomic systems. Due to the process of the design almost every second (44%) of the initially selected 465,768 variants could be designed and led to an array including 205,362 novel variants and 112,854 so called "routine variants" from previously applied arrays including common SNP markers considered for genomic selection. Subsequently, this custom array, called SWISScow array, was used routinely in the Swiss genomic breeding program. Therefore, comprehensive genotypes of 13,667 Swiss cattle are available, of which 5603 are Holstein.

Linkage disequilibrium (LD) (r^2) were estimated between diplotypes for haplotypes and both, custom array genotypes (a) and WGS-driven genotypes (b), in a merged dataset encompassing 5 603 (a) and 37 (b) animals, respectively. These analyses were conducted using plink v1.9[43].

For visualisation of the results from the different genome-wide analyses the R package OmicCircos was used[44]. Figure 1 summarizes the results from the haplotype identification using the trio and pgp approach, indicates if there are linked marker on the custom array, if the haplotypes show associations to current traits of the breeding program, and if there are significant GWAS associations in the Swiss Holstein population for the trait groups fertility, birth and survival.

In order to improve the interpretation of candidate variants conservation scores PhyloP and PhastCons from the UCSC database were taken into account[45,46]. Therefore, the tool LiftOver from UCSC tools was used to lift the bovine variants from the reference sequence ARS-UCD1.2[31] to the human genome 38[47]. For these positions, the conservation scores of 99 vertebrates could be extracted directly. Furthermore, for a more comprehensive prediction of the impact of protein-changing variants the tool PROVEAN[48] was applied. Lastly, the data provided by the 1000 Bull Genomes project run 8 was analysed for the distribution of candidate variants, as it is an international control cohort with a variety of breeds[41].

Results

Detection of novel and previously described haplotypes. All identified haplotypes are described in Table 3, their statistical evidence is summarized in Table S3 and the two outer circles in Fig. 1 demonstrate their distribution across the 29 bovine autosomes. The search for haplotypes deviating from HWE led to the detection of 24 haplotype regions, of which five had been previously described (Table 3, Table S3). These identified haplotypes were named in accordance to 17 previously described Holstein haplotypes (Table S1) and designated as subsequent haplotypes (HH18–HH38) (Table 3, Table S3). The obtained results from the trio-based approach led to the detection of 10 haplotypes, of which 6 were never observed in homozygous state, although at least 10 were expected (Table 3). On the other hand, the pgp-based approach revealed 21 haplotypes, of which 4 never occurred in homozygous state. Combining the outcome of both approaches revealed 7 commonly detected haplotype regions. The most significant deficient homozygous haplotype is the previously described haplotype HH21 for which 157 homozygous carriers were expected, but none were observed (Table 3, Table S3)[18]. The identified haplotypes have an average length of 1.12 Mb and range from 0.24 up to 2.56 Mb and segregate at an average allele frequency of 0.024 ranging from 0.014 to 5.46 (Table S3).

Association studies indicate phenotypic relevance of deficient homozygous haplotypes. Significant associations ($p < 0.05$) were detected for 21 haplotypes distributed across 17 chromosomes (Fig. 1, Table 3). Nevertheless, three of the detected deficient homozygous haplotypes show no significant associations to any of the 18 considered traits (Table S4). Most of the deficient homozygous haplotypes are associated with various fertility- and birth-related traits, whereas only two detected haplotypes show significant association to two traits of calf survival (Table 3). For example, HH37 shows significant associations to totally six different analysed fertility- and birth-related traits (Table 3). Detailed results from the association studies including all 24 detected deficient homozygous haplotypes and 18 traits of interest are summarised in Table S4. The by far strongest association was detected for the HH5 haplotype to the trait NRh (heifers non-return rate after 56 days) ($p = 0.00006$), which is in accordance with previously reported findings that a gross genomic deletion on chromosome 9 encompassing the entire *TFB1M* gene causes embryonic lethality[26]. Another example of a deficient homozygous haplotype showing a negative association to non-return rate in heifers is HH24 ($p = 0.02$), beside further significant associations of that haplotype to three birth-related traits (Table 3). The fertility-related trait

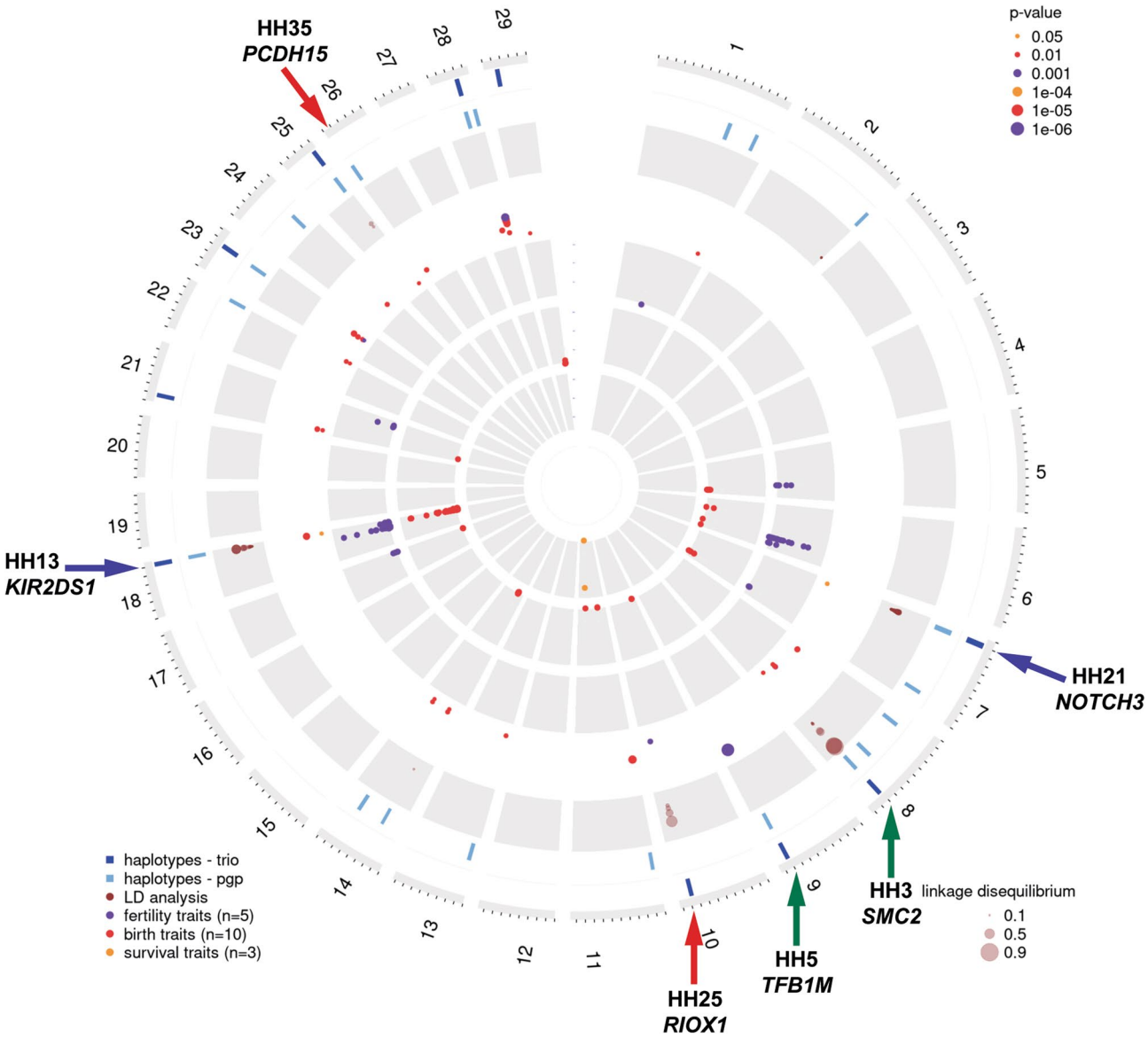


Figure 1. Summary of the SNP and WGS data analyses of the Swiss Holstein population. This plot visualises all the comprehensive genomic analysis. The most outer circles show the haplotypes per chromosome with reduced homozygosity for the trio approach in dark blue and the pgp approach in light blue. LD (r^2) between haplotypes and markers of the custom SNP array is indicated in the circle with the brown dots. Note the dot size correlates with the LD extent. The fourth circle shows significant results from the haplotype to phenotype association analyses. Note that the three evaluated trait groups are represented by different colours and the dot size correlates with the significance values. The three inner circles present the significant ($p < 0.0000043$) GWAS results across the fertility (purple), birth (red) and survival (yellow) traits. Scales are based on the $-\log10$(p-value). Note the green arrows indicating previously identified haplotypes and candidate causal variants. The blue arrows indicate the previously identified haplotypes HH21 and HH13[13,16,18] and the herein described candidate causative variants in the genes *NOTCH3* and *KIR2DS1,* respectively. Finally, the red arrows indicating the newly identified haplotypes HH25 and HH35 harbouring most likely causative variants in the genes *RIOX1* and *PCDH15*.

interval DFS (calving to first insemination) is negatively associated with the HH25 haplotype and the trait IFLc (interval fist to last insemination) with the HH32 haplotype whereas HH32 and HH30 show a reducing effect on the trait BWm (birth weight) (Table S4). The herein described deficient homozygous haplotype HH21 most likely corresponds to the previously described haplotypes 175.5 (VanRaden et al.[13]) and 07-126[18] and shows a negative association to the survival-related traits P2b (survival from day 3 up to 30 days) (Table 3). The also previously published deficient homozygous haplotype HH3 shows significant association to the birth-related trait GLm (gestation length) (Table 3), which can be explained with the already known *SMC2*-related abortion of homozygotes[24]. Finally, for the further two haplotypes with putative known causal variants (see below) HH35

Name[a]	chr	Pos in Mb[b]	Analysis	Homozygous haplotype carrier			Allele frequency %	Known and novel associated gene[c]	Associated traits		
				observed	expected	deficient			Fertility	Birth	Survival
HH18	1	10.302–104.376	pgp	7	28	75%	2.39				
HH19[d]	1	139.874–140.643	pgp	12	47	74%	3.00		GLd		
HH20	2	134.477–135.148	pgp	1	14	93%	1.66				
HH21[e]	7	7.872–10.433	trio/pgp	0	157	100%	5.46	*NOTCH3*			P2b
HH22	7	93.582–94.642	pgp	2	16	88%	1.73				
HH23	8	15.562–16.863	pgp	5	25	80%	2.20			MBm	
HH24	8	66.699–68.006	pgp	2	59	97%	3.36		NRh	BWm, CEm, MBm	
HH3[f]	8	90.959–92.085	trio/pgp	0	10	100%	1.38	*SMC2*	GLm		
HH5[g]	9	90.928–91.824	trio/pgp	0	30	100%	2.39	*TFB1M*	NRh		
HH25	10	86.876–87.773	trio	5	20	75%	1.96	*RIOX1*	DFS		
HH26	11	4.531–5.362	pgp	7	34	79%	2.56		GLd		
HH27	13	7.083–8.297	pgp	2	17	88%	1.82		GLm		
HH28	14	24.591–24.827	pgp	3	20	85%	1.98			SBd, SBm	
HH29	14	58.128–59.238	pgp	14	52	73%	3.14			SBm, CEm	
HH13[h]	18	60.932–62.101	trio	1	17	94%	1.81	*KIR2DS1*		BWd	P1
		62.070–63.045	pgp	2	22	91%	2.06				
HH30	21	7.844–8.671	trio	1	24	96%	2.13			BWd, CEd, CEm	
HH31	22	59.811–60.704	pgp	3	17	82%	1.80		GLd, GLm		
HH32	23	32.773–33.726	pgp	1	13	92%	1.57			BWd, CEm	
		33.545–34.948	trio	0	13	100%	1.55		GLd, IFLc	CEm	
HH33	24	44.693–45.678	pgp	14	42	67%	2.82			SBd	
HH34	25	36.162–37.330	trio /pgp	0	18	100%	1.86		GLm		
HH35	26	3.359–4.235	pgp	7	31	77%	2.44	*PCDH15*		SBm	
HH36	28	28.541–29.736	pgp	2	15	87%	1.71			CEm	
		29.731–30.877	trio	2	26	92%	2.25				
HH37	28	40.348–41.271	pgp	6	66	91%	3.54		GLd, GLm, IFLh	CEm, SBd, MBm	
HH38	29	14.243–15.317	trio	0	14	1.00	1.66			CEm	

Table 3. Haplotypes indicating reduced homozygosity in Swiss Holstein. *BWd* weight of calve at birth in kg, direct trait, *BWm* weight of calve at birth in kg, maternal trait, *CEd* direct calving ease, scored between 1-without help to 5-dystocia, *CEm* maternal calving ease, scored between 1-without help to 5-dystocia, *DFS* interval from calving to first service in days, *GLd* days from successful insemination to birth, direct trait, *GLm* days from successful insemination to birth, maternal trait, *IFLc* interval between first and last insemination for cows in days, *IFLh* interval between first and last insemination for heifer in days, *MBm* percentage of multiple births, *NRh* heifers non-return rate after 56 days, binary, *P1* survival from day 3 up to 30th day of life, *P2b* survival of young bulls from 31 up to 183 days, *SBd* percentage of calves born alive, direct trait, *SBm* percentage of calves born alive, maternal trait. [a]HH meaning Holstein haplotype. [b]According to the reference sequence ARS-UCD1.2[31]. [c]According to NCBI Annotation Release 106[91]. [d]Haplotype co-localises with previously described haplotype HHB by Shuster et al. and Cole et al.[51,85]. [e]Haplotype described before as 175.5 and 07–126 by VanRaden et al. and Sahana et al.[13,18]. [f]Haplotype previously described by VanRaden et al., McClure et al., Sahana et al. and Wu et al.[13,18,19,87]. [g]Haplotype previously described by Schütz et al. and Fritz et al.[26,27]. [h]Haplotype previously described by Fritz et al.[16].

shows a negative effect for the birth-related trait SBm (live births) while HH13 is associated to BWm (birth weight) as well as the survival related trait P1 (survival from day 3 up to 30th day of life) (Table 3).

In addition, the GWAS results are visualised using Manhattan plots in the supplementary material (Supplementary Figs. S1–S18). The genome-wide significant ($p < 0.0000043$) results are collected in Table S5 and in summary displayed in the three inner circles of Fig. 1. Among the five studied fertility-related traits we identified clear associations to a locus on chromosome 18 for the three traits DFS (calving to insemination), NRh (non-return rate for heifer) and IFLc (interval between first and last insemination for cows), that co-localize with the region of the herein detected deficient homozygous HH13 (Table S5). In addition, at the same locus interesting associations to five different birth-related traits were observed (Table S5). At the genome region of HH28 on chromosome 14, a GWAS hit to the two birth-related traits direct traits BWd (birth weight) and CEd (percentage normal births) was observed (Table S5). No co-localization between GWAS signals of the three studied survival traits and the herein detected deficient homozygous haplotypes could be identified. Nonetheless, a GWAS hit on chromosome 11 for the calf survival-related trait P2h (survival of heifers from 31 up to 458 days) was observed at the identical genome region on chromosome 11 of the previously described CDH haplotype containing a deleterious *APOB* loss-of-function variant causing rearing loss (Table S1)[29].

Haplotype	Gene	OMIM	Variant description				
			Genomic position[a]	Description	Transcript[b]	Coding DNA change	Protein change
HH21[c]	*NOTCH3*	600276	chr7:7913459	3 bp deletion (inframe deletion)	XM_003586246.3	c.129_131delTTG	p.Cys44del
HH3[d]	*SMC2*[e]	605576	chr8:93753358	SNV (missense)	XM_015472668.2	c.3404T > C	p.Phe1135Ser
HH25	*RIOX1*	611919	chr10:84938370	30 bp deletion (inframe deletion)	NM_001099702.1	c.396_425delGGCGCAGACCCC GGCGGCACGCTTGGTGGA	p.Ala133_Glu142del
HH13[f]	*KIR2DS1*	604952	chr18:62758881	SNV (nonsense)	NM_001097567.1	c.475C > T	p.Gln159*
HH35	*PCDH15*	605514	chr26:5325675	SNV (missense)	XM_015460562.2	c.2599C > G	p.Leu867Val

Table 4. Candidate variants potentially explaining the identified haplotype regions. [a]According to the reference sequence ARS-UCD1.2[31]. [b]According to NCBI Annotation Release 106[91]. [c]Haplotype described before as 175.5 and 07-126 by VanRaden et al. and Sahana et al., respectively[13,18]. [d]Haplotype previously described by VanRaden et al., McClure et al., Sahana et al. and Wu et al.[13,18,19,87]. [e]Variant previously detected by McClure et al.[24]. [f]Haplotype previously described by Fritz et al.[16].

Identification of four novel potentially causal variants. The generated comprehensive WGS data was mined in genome regions containing deficient homozygous haplotype for variants potentially causing pre- or post-natal lethality or sublethal conditions. By filtering for variants indicating an obvious depletion in homozygosity and showing a strong LD with the haplotype, we successfully uncovered five protein-changing DNA variants as candidates (Table 4). This short list of possibly causal variants includes the previously described missense mutation in the *SMC2* gene associated with HH3-related embryonic lethality[24] (Table S6). In addition, we detected four novel variants linked to different deficient homozygous haplotypes affecting these four genes: *NOTCH3* (HH21), *RIOX1* (HH25), *KIR2DS1* (HH13), and *PCDH15* (HH35) (Table 4). Unfortunately, only two of these five variants, those affecting *SMC2* and *KIR2DS1*, could be successfully added to the customized SWISScow array and thereby genotyped in several thousand animals of the current Swiss dairy population. In accordance to the results of the haplotype analysis for HH3 and HH13, no single homozygous carrier of the *SMC2* and *KIR2DS1* variants was observed neither in the current population of more than 14 thousand Swiss dairy cattle nor in any other breed of cattle included in the 1000 Bull Genomes project (Table S6).

As reported earlier by Ref.[24] the *SMC2* missense variant p.Phe1135Ser affects an evolutionary conserved residue of the C-terminal P-loop-nucleoside triphosphate hydrolase domain classified as a deleterious change according to in silico predictions. In the studied Swiss Holstein population this variant occurs a low frequency (1.3%) in high LD ($r^2 = 0.85$) with the HH3 haplotype on chromosome 8 at ~ 94 Mb (Table S6). The *SMC2* variant was absent from all other studied breeds (Table S6).

The detected stop-gain variant in the *killer cell immunoglobulin like receptor, two Ig domains and short cytoplasmic tail 1* (*KIR2DS1*) gene showing missing homozygosity is located on chromosome 18 at ~ 62 Mb in LD ($r^2 = 0.49$) with the HH13 haplotype. The detected *KIR2DS1* variant located in exon 4 leads to the premature termination of the protein sequence at codon 159 (p.Gln159*) truncating the protein by almost two third including functionally important domains (Fig. 2A). It segregates within the Swiss Holstein population at an allele frequency of 5% and shows a highly significant deviation from HWE ($p = 0.0000007$) based on a Chi-square test (Table S6). Interestingly the *KIR2DS1* variant obviously occurs in various other breeds beside Holstein cattle (Table S6).

Further, we propose three coding variants in *NOTCH3*, *RIOX1,* and *PCDH15* as candidate causal variants that were so far not evaluated in the broader population due to limitations in the design process of the customized SWISScow array.

Firstly, we identified a disruptive in-frame deletion in the *notch receptor 3* (*NOTCH3*) gene moderately linked ($r^2 = 0.49$) with the completely deficient homozygous haplotype HH21 on chromosome 7 at ~ 9 Mb (Table 4). This protein-changing variant knocks-out a highly conserved cysteine residue (p.Cys44del) of an epidermal growth factor-like repeats (EGFRs) that comprise the extracellular domain of the NOTCH3 receptor. The deletion results in a loss of a cysteine residue in the first of these numerous EGFRs (Fig. 2B) and is predicted to be deleterious (Table S6). This variant occurs at a low allele frequency of 0.014 and was never observed in homozygous state in any of the more than 4000 bovine genomes with available WGS data (Table S6). The studied *NOTCH3* variant appears to be very rare, and among the 4109 animals of the 1000 Bull Genomes project, it was found in only three animals of different breeds (Table S6).

Secondly, another disruptive in-frame deletion encompassing 30 nucleotides of the open reading frame of the *ribosomal oxygenase 1* (*RIOX1*) gene was detected that occurs perfectly linked ($r^2 = 1$) with the HH25 haplotype on chromosome 10 at ~ 87 Mb (Table S6). Interestingly, within our complete WGS data a single homozygous animal was detected, whereas 16 homozygotes were present in the 1000 Bull Genomes variant catalogue. By visual inspection of the mapping quality of the reads around this variant, we noticed that the mapping algorithm (Supplementary Fig. S19) did not splice many of the reads covering the 30 bp deletion properly. Therefore, we assume that these homozygotes most likely represent false positives and are actually only heterozygous carriers of the *RIOX1* variant. Obviously, the *RIOX1* deletion occurs in many breeds other than Holstein (Table S6). The disruptive in-frame deletion affects the N-terminal region of the RIOX1 protein (p.Ala133_Glu142del) (Fig. 2C).

Lastly, we identified a missense variant in the *protocadherin related 15* (*PCDH15*) gene that is perfectly linked ($r^2 = 1$) with the HH35 haplotype on chromosome 26 at ~ 4 Mb (Table S6). Interestingly, for this variant, as well as for the HH35 haplotype we observed some homozygous carriers (Table 3, Table S6). In the 1000 Bull Genomes

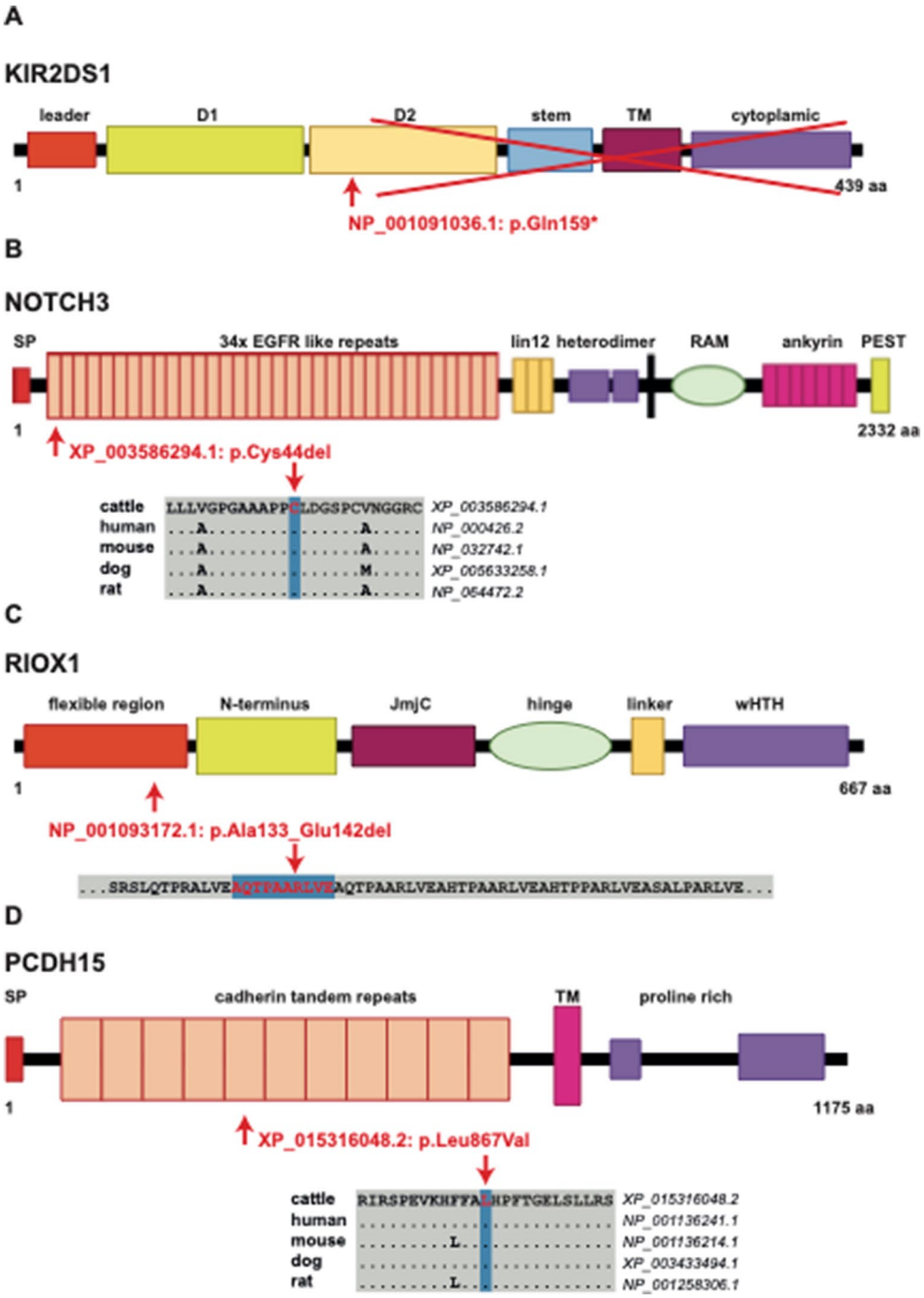


Figure 2. Features of the four novel candidate causal variants. Note the predicted protein changes of the detected variants indicated in red. ((**A**) *KIR2DS1*:p.Gln159*[82], (**B**) *NOTCH3*:p.Cys44del[76], (**C**) *RIOX1*:p.Ala133_Glu142del[83], (**D**) *PCDH15*:p.Leu867Val[84]).

Project variant catalogue all identified variant carriers were indicated as Holstein, Ayrshire and Norwegian Red (Table S6). Nevertheless, the amino acid exchange (p.Leu867Val) affects a highly conserved residue of a cadherin tandem repeat domain (Fig. 2D) although the computer predicted consequence was inconclusive (Table S6).

Discussion

After giving an overview of the known haplotypes with missing homozygosity in the worldwide Holstein breed, we have conducted a comprehensive survey of genomic and phenotypic data of the Swiss Holstein cattle population to search for further deficient homozygous haplotypes. We focussed on the detection of hidden recessive variants and found five candidates that cause the observed missing homozygosity. These variants can lead to the exclusion from the population due to natural selection processes, e.g., embryonic lethality, abortions and

stillbirth, or due to artificial selection against animals that show illthrift or other non-lethal more subtle conditions, e.g., metabolic disorders that are often not diagnosed.

In previous studies performed in Holstein cattle a simple allele frequency-based approach was applied to identify haplotypes showing a depletion of homozygous animals by applying the assumption of random mating and a Chi-squared test[13,18,49]. Other studies accounted for the mating status by included the haplotype state of the sire and the maternal grandsire[16,50]. In comparison, the herein presented study uses two different approaches, a trio- and pgp-approach, that include genotyped trios and applies a Fisher exact test. This had the advantage that we were able to detect haplotypes that segregate at a lower allele frequency. The pgp-based approach includes almost double the data from the trio-based approach. This is explained by the genotyping management, in which male breeding animals are usually genotyped, but by far not all females. Therefore, we identified almost twice the number of haplotypes with the pgp-approach. Nevertheless, the applied trio-approach is assumed to be more powerful as the parental haplotype inheritance can be traced directly.

Introducing this novel approach, we were able to verify our outcomes with the results of earlier studies in different Holstein cattle populations. With the applied methods, we were able to detect four haplotype regions, namely HH3, HH5, HH13 and HH21 that were previously described to occur in several subpopulations of Holstein cattle[13,16,18,26,51]. Another herein detected haplotype (HH19) co-localizes with the genome regions of the known deficient homozygous haplotype HHB on the proximal end of chromosome 1. However, it obviously differs to the BLAD-associated HHB haplotype as we confirmed by using the customized SWISScow array that the causative *ITGB2* variant does not segregate anymore in the Swiss Holstein population (results not shown) probably due to the strict exclusion of carriers in the early 1990s.

Furthermore, by applying the WGS data screen, we confirmed the known Holstein breed-specific HH3-linked missense variant in *SMC2* on chromosome 8 that also in our data showed complete missing homozygosity[24]. Interestingly, for this deleterious variant we could only detect a single significant phenotypic effect in our additive models for gestation length based on the haplotype, but a suggestive negative effect on non-return rate in heifer. Negative effects on the number of inseminations in heifers within that genomic region were shown in Canadian Holstein cattle[52]. Nevertheless, these findings support a purely recessive embryonic lethal effect of the *SMC2* variant that allows merely the indirect detection in non-return rate[24]. For HH5, we were unfortunately unable to confirm the association with the most likely causative gross deletion encompassing the *TFB1M* gene as we focused in the scope of this project on SNV and short indel genotyping.

Interestingly, we found a co-localization of a GWAS peak for survival during the first 458 days of life in the region of the previously reported CDH haplotype[29]. The CDH haplotype was shown to be linked with a pathogenic endogenous retroviral insertion into the *APOB* gene and thereby leading to a subvital malabsorption of cholesterol, provoking increased juvenile mortality[26,29,30,53]. This haplotype was difficult to detect due to it segregating in homozygous state in the adult population, as the LD between the CDH haplotype and the *APOB* variant was not perfect as the ancestral version of the haplotype still occurred[29,30]. Initially, cholesterol deficiency was thought to be a recessive or codominant inherited disorder, what led to the identification of the responsible locus by applying a case–control based GWAS[29]. Further studies examining heterozygous carriers showed that the *APOB*-associated disorder is not a simple Mendelian disease, as heterozygous animals can also show severe clinical signs of cholesterol deficiency[54]. The identified GWAS peak in the current study underlines these findings with an additive model. For a monogenic recessive lethal disorder, one would not expect to see any effect within an additive model, as all the homozygous affected animals would be more or less excluded from the genotyped population and it can be assumed that these heterozygous animals are phenotypically normal like the homozygotes. It is in line with recent findings in cattle that additive models are shown to have their limitations and non-additive models should be taken into account for further phenotypic evaluations of recessive disorders[55].

We propose to add four novel candidate causative variants to the list of Holstein haplotypes with known cause for missing homozygosity. We emphasize to highlight two variants in *KIR2DS1* and *NOTCH3* linked to the abovementioned confirmed deficient homozygous haplotypes HH13 and HH21, respectively. Both detected variants segregate in Holstein cattle but, in contrast to the HH3-linked *SMC2* variant, represent derived alleles that predates establishing of modern breed as they rarely occur also in unrelated breeds of cattle.

Interestingly, in human *KIR2DS1* is reported to be associated with reproduction, namely placentation success[56,57]. *KIR2DS1* is an activator receptor, belongs to the KIR family of natural killer cell Ig-like receptors and plays a vital role in the immune system (OMIM #604952). Regarding the involvement in disorders, *KIR2DS1* is known to be associated in human with an autoimmune disease called psoriasis vulgaris[58,59] and leukaemia[60–62]. In later studies reproductive failure and foetal growth restriction were associated with maternal KIR2DS1 and foetal leukocyte antigen C2 (HLA-C2) imbalance, however, being complicated by the interference of the maternal and the embryonic genotype[56,63–66]. Nonetheless, the function of KIR2DS1 was described as protective against pregnancy loss[66,67]. Moreover a study showed the importance of KIR2DS1/HLA-C2 in response to viral placental infections[68]. The herein detected bovine variant most likely represents a true loss-of-function mutation of *KIR2DS1*, as possibly nonsense-mediated decay selectively recognizes and degrades mRNAs whose open reading frame is truncated by a premature translation termination codon. Therefore, this variant is supposed to lead to pregnancy losses of homozygous embryos. Furthermore, as the function of KIR2DS1 is sensitive to its expression[68], we speculate that also a negative effect in heterozygous carriers might exist. This hypothesis is supported by GWAS hits in the genome region of HH13 for several fertility-related traits (interval calving to insemination, non-return rate heifer and interval first and last insemination), as well as birth-related traits (percentage normal births, percentage live births, birth weight as well as gestation length for maternal and direct traits). Interestingly, major QTLs for fertility and birth traits were mapped at 60 Mb on chromosome 18 in the German Holstein population[69].

The *NOTCH3* gene belongs to the mammalian Notch family that includes four proteins, which are transmembrane developmental signalling receptors[70–72]. The Notch signalling pathway has important roles in the

development of organs, regulation of cell death, cell survival and cell abundance[70–72]. *NOTCH3* is known to be associated with fatal dominant inherited disorders such as myofibromatosis, cerebral arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) and lateral meningocele syndrome (OMIM # 600276), as well as different forms of cancer[73]. In *Notch3*-deficient mice, it was initially shown that *Notch3* seems to be non-essential for proper embryonic development and fertility[74], however, post-natal pathology in *Notch3*-null mice revealed structural defects in the smooth muscle cells, arterial differentiation and the arterial morphology[75]. The *NOTCH3* variant linked to the HH21 haplotype is predicted to delete a highly conserved cysteine residue in the first of 34 EGFR modules that are characterised by six highly conserved cysteine's that are essential for the stability of each EGFR-module fold by their three disulphide bonds[76]. EGFR1–3 had been shown to be critical for Notch pathway activation[77]. Interestingly, the most significantly associated trait with HH21 was survival of young bulls from 31 up to 183 days, which indicates a possibly subvital condition caused by the *NOTCH3* variant. So far, the proposed variant was never observed in homozygous state in any of more than 4000 cattle genomes, while a population evaluation in Swiss Holstein is still pending to prove if there are indeed 100% no homozygous carriers as to be assumed from the haplotype analysis.

Lastly, we would like to propose two candidate causal variants in *RIOX1*, and *PCDH15*, that show perfect LD to the deficient homozygous haplotypes HH25 and HH35, respectively. *RIOX1* is an oxygenase that can act as both a histone lysine demethylase and a ribosomal histidine hydroxylase and plays a central role in histone code, chromatin organisation and DNA transcription regulation (OMIM #611919). Due to the basic importance of these processes for DNA and cell function, we speculate that a disruption of the protein by a motif of 10 amino acids might lead to disturbance of the meiosis or to the inactivation of cell differentiation in very early stages of the embryogenesis. This hypothesis would be supported by the haplotype association for HH25 that indicates a prolonged interval calving to first insemination (DFS). Here we also would like to address the fact that such small deletions affecting repetitive DNA motifs clearly show up the limitations of the applied short-read methods for whole-genome sequencing and subsequent indel calling (Supplementary Fig. S19). Nevertheless, because the variant *RIOX1* allele obviously occurs in a variety of breeds, this variant might represent a quite old mutation that arose already before modern breed formation that could also be of less impact. If the proposed variant is indeed reduced in number of homozygous carriers needs to be evaluated by further genotyping.

Finally, we propose a missense variant in *PCDH15* as strong causal candidate for the depletion in homozygosity of HH35. The occurrence of this bovine variant is limited to Holstein and animals of populations with documented introgression of Holstein cattle. It affects an evolutionary conserved residue of a functionally important domain and the affected gene is a member of the cadherin superfamily[78]. It encodes an integral membrane protein that is known to be vital for the maintenance of normal retinal and cochlear function (OMIM #605514). In men, the affected gene *PCDH15* is associated with recessively inherited forms of deafness and neurosensorial Usher syndrome[79]. The HH35 haplotype is not completely deficient homozygous indicating no embryonic lethal effect of the associated *PCDH15* variant. Therefore, we speculate that possibly a subtle disorder leads to the social exclusion of young individuals, probably due to their abnormal behaviour as they have a different perception of the environment than unaffected animals. Alternatively, we hypothesize that possibly not all biological functions of PCDH15 are known and more severe conditions could arise leading to pre- or neonatal lethality and thereby explaining the lack of homozygotes. Nevertheless, we suggest also for this newly described variant in *PCDH15* to follow up by extended genotyping and propose a detailed clinical examination of homozygous living animals to evaluate their individual health status.

At this point we would like to lay out some limitations of the presented study. The imputation step, which is based on pedigree information and allele frequencies, can affect the detection of decreased homozygosity. Therefore, imputation has limited power for recently arisen variants due to the existence of ancestral and derived versions of identical haplotypes. This could potentially lead to the false negative detection of recent haplotype regions and the false positive detection of haplotype regions due to naturally over-represented alleles. Furthermore, the accuracy of imputation depends on the SNP density of the genotype array and number of genotypes available per array[80]. During the design of the custom genotyping array, it was not possible to include all variants of interest. Mainly due to technical reasons, but also due to the more complicated types of variants (e.g. deletions). The manufacturer did not recommend those variants since in silico predicted genotyping quality was expected to be limited. Nevertheless, as shown for the HH3 and HH13 associated variants in *SMC2* and *KIR2DS1*, the custom array was a very efficient tool to monitor several thousands of animals. The displayed results illustrate that even though Holstein is an international breed local selection strategies affect the genomic structure of its subpopulations. From a genome-wide perspective, it might be possible that we have overlooked additional haplotype regions due to the restricted sliding-window length of 50 SNPs, considering that there is a correlation between window size and the detected homozygosity[36]. In addition, the restriction for protein-changing SNVs and indels limits the identification for candidate causal variants. We have not considered non-coding regulatory variants or larger structural variants that have a pathogenic impact, which we do not know yet. An example for this is the HH5 associated large deletion including the gene *TFB1M*[26], for which we could confirm the haplotype region in our data. Furthermore, it was shown that graph based reference sequences can improve the annotation and could potentially lead to the detection of further protein-changing variants[81]. In order to improve the understanding of the described candidate variants, further studies are recommended to evaluate all homozygous haplotype and variant carriers. As the example of CDH in Holstein cattle has taught, the oversimplification of the depletion of homozygosity by assuming the cause to be a monogenetic recessive disorder might be misleading.

In conclusion, we demonstrate that with mining massive SNP data in combination with WGS data, we mapped 19 novel and 5 previously described haplotype regions and unravel novel recessive protein-changing variants in *NOTCH3*, *RIOX1*, *KIR2DS1* and *PCDH15* segregating at low to moderate frequencies. In addition, our findings confirm previously identified loci such as *SMC2* for HH3. The phenotypic associations support additive effects in the described haplotype regions and variants. Taken together this study expands the spectrum of undesired alleles

impairing reproduction success in Holstein cattle, the world's most important dairy breed, enabling improved DNA-based selection in order to reduce reproductive failure and animal loss.

Data availability

The whole-genome sequencing data is stored in the European Nucleotide Archive and can be found with the sample IDs available in Table S2. For access to the SNP genotyping data, we ask interested people to contact the authors or the breeding associations directly.

Received: 4 October 2021; Accepted: 7 March 2022
Published online: 31 March 2022

References

1. Hansen, L. B. Consequences of selection for milk yield from a Geneticist's viewpoint. *J. Dairy Sci.* **83**, 1145–1150 (2000).
2. Zenger, K. R., Khatkar, M. S., Cavanagh, J. A. L., Hawken, R. J. & Raadsma, H. W. Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. *Anim. Genet.* **38**, 7–14 (2007).
3. Ma, L. *et al.* Genome changes due to artificial selection in U.S. Holstein cattle. *BMC Genomics* **20**, 1–14 (2019).
4. Makanjuola, B. O. *et al.* Effect of genomic selection on rate of inbreeding and coancestry and effective population size of Holstein and Jersey cattle populations. *J. Dairy Sci.* **103**, 5183–5199 (2020).
5. Pryce, J. E. *et al.* World trends in dairy cow fertility. In *World Congress of Genetics Applied to Livestock Production*, Vol. 10, 6 (2014).
6. Walsh, S. W., Williams, E. J. & Evans, A. C. O. A review of the causes of poor fertility in high milk producing dairy cows. *Anim. Reprod. Sci.* **123**, 127–138 (2011).
7. Seegers, H., Beaudeau, F., Fourichon, C. & Bareille, N. Reasons for culling in French Holstein cows. *Prev. Vet. Med.* **36**, 257–271 (1998).
8. Rilanto, T. *et al.* Culling reasons and risk factors in Estonian dairy cows. *BMC Vet. Res.* **16**, 173 (2020).
9. Mao, X. *et al.* Genome-wide association mapping for dominance effects in female fertility using real and simulated data from Danish Holstein cattle. *Sci. Rep.* **10**, 1–9 (2020).
10. Cai, Z., Guldbrandtsen, B., Lund, M. S. & Sahana, G. Prioritizing candidate genes for fertility in dairy cows using gene-based analysis, functional annotation and differential gene expression. *BMC Genomics* **20**, 1–9 (2019).
11. Jiang, J. *et al.* A large-scale genome-wide association study in U.S. Holstein cattle. *Front. Genet.* https://doi.org/10.3389/fgene.2019.00412 (2019).
12. Minozzi, G. *et al.* Genome wide analysis of fertility and production traits in Italian Holstein cattle. *PLoS ONE* **8**, 1–10 (2013).
13. VanRaden, P. M., Olson, K. M., Null, D. J. & Hutchison, J. L. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* **94**, 6153–6161 (2011).
14. Puckowska, P., Borowska, A., Szwaczkowski, T., Oleński, K. & Kamiński, S. Effects of a novel missense polymorphism within the SIGLEC5 gene on fertility traits in Holstein–Friesian cattle. *Reprod. Domest. Anim.* **54**, 1163–1168 (2019).
15. Taylor, J., Schnabel, R. & Sutovsky, P. Review: Genomics of bull fertility. *Animal* **12**, s172–s183 (2018).
16. Fritz, S. *et al.* Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS ONE* **8**, e65550 (2013).
17. Hozé, C. *et al.* A splice site mutation in CENPU is associated with recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **103**, 607–612 (2020).
18. Sahana, G., Nielsen, U. S., Aamand, G. P., Lund, M. S. & Guldbrandtsen, B. Novel harmful recessive haplotypes identified for fertility traits in Nordic Holstein cattle. *PLoS ONE* **8**, 8–12 (2013).
19. Wu, X., Mesbah-Uddin, M., Guldbrandtsen, B., Lund, M. S. & Sahana, G. Haplotypes responsible for early embryonic lethality detected in Nordic Holsteins. *J. Dairy Sci.* **102**, 11116–11123 (2019).
20. Charlier, C. *et al.* A deletion in the bovine FANCI gene compromises fertility by causing fetal death and brachyspina. *PLoS ONE* **7**, e43085 (2012).
21. Thomsen, B. *et al.* A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. *Genome Res.* **16**, 97–105 (2006).
22. Adams, H. A. *et al.* Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in Holstein dairy cattle. *J. Dairy Sci.* **99**, 6693–6701 (2016).
23. Daetwyler, H. D. *et al.* Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* **46**, 858–865 (2014).
24. McClure, M. C. *et al.* Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative Causative mutation in SMC2 for HH3. *PLoS ONE* **9**, e92769 (2014).
25. Cooper, T. A., Wiggans, G. R., Null, D. J., Hutchison, J. L. & Cole, J. B. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. *J. Dairy Sci.* **97**, 3878–3882 (2014).
26. Schütz, E. *et al.* The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. *PLoS ONE* **11**, e0154602 (2016).
27. Fritz, S. *et al.* An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **101**, 6220–6231 (2018).
28. Ortega, M. S. *et al.* Truncation of IFT80 causes early embryonic loss in cattle. *BioRxiv* **2021**, 99. https://doi.org/10.1101/2021.07.02.450952 (2021).
29. Kipp, S. *et al.* Identification of a haplotype associated with cholesterol deficiency and increased juvenile mortality in Holstein cattle. *J. Dairy Sci.* **99**, 8915–8931 (2016).
30. Menzi, F. *et al.* A transposable element insertion in APOB causes cholesterol deficiency in Holstein cattle. *Anim. Genet.* **47**, 253–257 (2016).
31. National Center for Biotechnology Information. *ARS-UCD1.2 Assembly* (National Center for Biotechnology Information, 2018).
32. Rosen, B. D. *et al.* De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **9**, 021 (2020).
33. Sargolzaei, M., Chesnais, J. P. & Schenkel, F. S. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* **15**, 478 (2014).
34. Sargolzaei, M. *SNP1101 User's Guide. Version 1.0* (Guelph HiggsGene Solution Inc., 2014).
35. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188 (2001).
36. Hoff, J. L., Decker, J. E., Schnabel, R. D. & Taylor, J. F. Candidate lethal haplotypes and causal mutations in Angus cattle. *BMC Genomics* **18**, 799 (2017).
37. Garrick, D. J., Taylor, J. F. & Fernando, R. L. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* **41**, 55 (2009).

38. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
39. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
40. Häfliger, I. M. *et al.* Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **51**, 382–390 (2020).
41. Hayes, B. J. & Daetwyler, H. D. 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: Applications and outcomes. *Annu. Rev. Anim. Biosci.* **7**, 89–102 (2019).
42. Brentp. goleft. *Github* (2018). https://github.com/brentp/goleft. Accessed 4 Oct 2021.
43. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
44. Hu, Y. *et al.* OmicCircos: A simple-to-use R package for the circular visualization of multidimensional Omics data. *Cancer Inform.* **13**, 13–20 (2014).
45. Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
46. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
47. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
48. Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* **7**, e46688 (2012).
49. Wu, X., Mesbah-Uddin, M., Guldbrandtsen, B., Lund, M. S. & Sahana, G. Novel haplotypes responsible for prenatal death in Nordic Red and Danish Jersey cattle. *J. Dairy Sci.* **103**, 4570–4578 (2020).
50. Derks, M. F. L. *et al.* A systematic survey to identify lethal recessive variation in highly managed pig populations. *BMC Genomics* **18**, 858 (2017).
51. Shuster, D. E., Kehrli, M. E. Jr., Ackermann, M. R. & Gilbert, R. O. Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9225–9229 (1992).
52. Makanjuola, B. O. *et al.* Identification of unique ROH regions with unfavorable effects on production and fertility traits in Canadian Holsteins. *Genet. Sel. Evol.* **53**, 68 (2021).
53. Artesi, M. *et al.* Pooled CRISPR Inverse PCR sequencing (PCIP-seq): Simultaneous sequencing of retroviral insertion points and the integrated provirus with long reads. *BioRxiv* **2019**, 558130 (2019).
54. Häfliger, I. M. *et al.* APOB-associated cholesterol deficiency in Holstein cattle is not a simple recessive disease. *Anim. Genet.* **50**, 372–375 (2019).
55. Reynolds, E. G. M. *et al.* Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. *Nat. Genet.* **53**, 949–954 (2021).
56. Xiong, S. *et al.* Maternal uterine NK cell-activating receptor KIR2DS1 enhances placentation. *J. Clin. Investig.* **123**, 4264–4272 (2013).
57. Faridi, R. M. & Agrawal, S. Killer immunoglobulin-like receptors (KIRs) and HLA-C allorecognition patterns implicative of dominant activation of natural killer cells contribute to recurrent miscarriages. *Hum. Reprod.* **26**, 491–497 (2011).
58. Łuszczek, W. *et al.* Gene for the activating natural killer cell receptor, KIR2DS1, is associated with susceptibility to psoriasis vulgaris. *Hum. Immunol.* **65**, 758–766 (2004).
59. Suzuki, Y. *et al.* Genetic polymorphisms of killer cell immunoglobulin-like receptors are associated with susceptibility to *Psoriasis vulgaris*. *J. Investig. Dermatol.* **122**, 1133–1136 (2004).
60. Marin, D. *et al.* KIR2DS1 genotype predicts for complete cytogenetic response and survival in newly diagnosed chronic myeloid leukemia patients treated with imatinib. *Leukemia* **26**, 296–302 (2012).
61. La Nasa, G. *et al.* Homozygosity for killer immunoglobin-like receptor haplotype A predicts complete molecular response to treatment with tyrosine kinase inhibitors in chronic myeloid leukemia patients. *Exp. Hematol.* **41**, 424–431 (2013).
62. Naumova, E. *et al.* Genetic polymorphism of NK receptors and their ligands in melanoma patients: Prevalence of inhibitory over activating signals. *Cancer Immunol. Immunother.* **54**, 172–178 (2005).
63. Dambaeva, S. V. *et al.* Recurrent pregnancy loss in women with killer cell immunoglobulin-like receptor KIR2DS1 is associated with an increased HLA-C2 allelic frequency. *Am. J. Reprod. Immunol.* **75**, 94–103 (2016).
64. Colucci, F. The role of KIR and HLA interactions in pregnancy complications. *Immunogenetics* **69**, 557–565 (2017).
65. Yang, X. & Meng, T. Killer-cell immunoglobulin-like receptor/human leukocyte antigen-C combination and 'great obstetrical syndromes' (review). *Exp. Ther. Med.* **22**, 1178 (2021).
66. Hiby, S. E. *et al.* Maternal activating KIRs protect against human reproductive failure mediated by fetal HLA-C2. *J. Clin. Investig.* **120**, 4102–4110 (2010).
67. Parham, P. & Guethlein, L. A. Pregnancy immunogenetics: NK cell education in the womb? *J. Clin. Investig.* **120**, 3801–3804 (2010).
68. Crespo, Â. C., Strominger, J. L. & Tilburgs, T. Expression of KIR2DS1 by decidual natural killer cells increases their ability to control placental HCMV infection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15072–15077 (2016).
69. Müller, M. P. *et al.* Genome-wide mapping of 10 calving and fertility traits in Holstein dairy cattle with special regard to chromosome 18. *J. Dairy Sci.* **100**, 1987–2006 (2017).
70. Kopan, R. & Ilagan, M. X. G. The canonical Notch signaling pathway: Unfolding the activation mechanism. *Cell* **137**, 216–233 (2009).
71. Bray, S. J. Notch signalling in context. *Nat. Rev. Mol. Cell Biol.* **17**, 722–735 (2016).
72. Baron, M. Combining genetic and biophysical approaches to probe the structure and function relationships of the notch receptor. *Mol. Membr. Biol.* **34**, 33–49 (2017).
73. Bellavia, D. *et al.* Notch3: From subtle structural differences to functional diversity. *Oncogene* **27**, 5092–5098 (2008).
74. Krebs, L. T. *et al.* Characterization of Notch3-deficient mice: Normal embryonic development and absence of genetic interactions with a Notch1 mutation. *Genesis* **37**, 139–143 (2003).
75. Domenga, V. *et al.* Notch3 is required for arterial identity and maturation of vascular smooth muscle cells. *Genes Dev.* **18**, 2730–2735 (2004).
76. Hosseini-Alghaderi, S. & Baron, M. Notch3 in development, health and disease. *Biomolecules* **10**, 485 (2020).
77. Kershaw, N. J. *et al.* Notch ligand delta-like1: X-ray crystal structure and binding affinity. *Biochem. J.* **468**, 159–166 (2015).
78. Pepermans, E. & Petit, C. The tip-link molecular complex of the auditory mechano-electrical transduction machinery. *Hear. Res.* **330**, 10–17 (2015).
79. Keats, B. J. B. & Savas, S. Genetic heterogeneity in Usher syndrome. *Am. J. Med. Genet. A* **130**, 13–16 (2004).
80. VanRaden, P. M. *et al.* Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* **96**, 668–678 (2013).
81. Crysnanto, D., Leonard, A. S., Fang, Z.-H. & Pausch, H. Novel functional sequences uncovered through a bovine multi-assembly graph. *Proc. Natl. Acad. Sci.* **118**, e2101056118 (2021).
82. Robinson, J., Waller, M. J., Stoehr, P. & Marsh, S. G. E. IPD—The immuno polymorphism database. *Nucleic Acids Res.* **33**, D523–D526 (2005).
83. Chowdhury, R. *et al.* Ribosomal oxygenases are structurally conserved from prokaryotes to humans. *Nature* **510**, 422–426 (2014).

84. Rouget-Quermalet, V. *et al.* Protocadherin 15 (PCDH15): A new secreted isoform and a potential marker for NK/T cell lymphomas. *Oncogene* **25**, 2807–2811 (2006).
85. Cole, J. B. *et al.* Haplotype tests for recessive disorders that affect fertility and other traits. *Animal Improvement Program Research Report* (2018). https://www.aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html. Accessed 4 Oct 2021.
86. Agerholm, J. S., McEvoy, F. & Arnbjerg, J. Brachyspina syndrome in a Holstein calf. *J. Vet. Diagnostic Investig.* **18**, 418–422 (2006).
87. McClure, M. *et al.* Fine mapping for weaver syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. *PLoS ONE* **8**, e59251 (2013).
88. Agerholm, J. S., Bendixen, C., Andersen, O. & Arnbjerg, J. Complex vertebral malformation in Holstein calves. *J. Vet. Diagnostic Investig.* **13**, 283–289 (2001).
89. Shanks, R. D., Dombrowski, D. B., Harpestad, G. W. & Robinson, J. L. Inheritance of UMP synthase in dairy cattle. *J. Hered.* **75**, 337–340 (1984).
90. Schwenger, B., Schöber, S. & Simon, D. DUMPS cattle carry a point mutation in the uridine monophosphate synthase gene. *Genomics* **16**, 241–244 (1993).
91. National Center for Biotechnology Information. *NCBI Bos Taurus Annotation Release 106* (National Center for Biotechnology Information, 2018).

Acknowledgements

The authors are grateful to the Swiss cattle breeding associations (swissherdbook, Holstein Switzerland) for providing genomic and phenotypic data. We thank Nathalie Besuchet-Schmutz for expert technical assistance. We thank the Interfaculty Bioinformatics Unit of the University of Bern for providing high performance computing infrastructure. In addition, we acknowledge Christy Vander Jagt, Amanda Chamberlain and Hans Daetwyler of the 1000 Bull Genomes Project for providing the variant catalogue. The Swiss National Science Foundation, Grant Number 172911, funds this study.

Author contributions

I.M.H. performed the whole-genome sequencing data preparation and analysis. Further, I.M.H. interpreted all collective results, drafted this manuscript with all tables and figures. C.D. acquired the whole-genome sequencing data and drafted this manuscript. M.S. and F.R.S. performed the SNP genotyping data preparation and analysis, including the GWAS and haplotype associations. C.D. and F.R.S. designed and managed the project. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-09403-6.

Correspondence and requests for materials should be addressed to I.M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reverse genetic screen for deleterious recessive variants in the local Simmental cattle population of Switzerland

Journal:	Animals
Manuscript status:	published
Contributions:	Whole-genome sequencing data preparation and analyzes, investigation, methodology, visualizations, writing original draft and revisions
Displayed version:	published version
DOI:	10.3390/ani11123535

animals

MDPI

Article

Reverse Genetic Screen for Deleterious Recessive Variants in the Local Simmental Cattle Population of Switzerland

Irene M. Häfliger [1], Franz R. Seefried [2] and Cord Drögemüller [1,*]

[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; irene.haefliger@vetsuisse.unibe.ch
[2] Qualitas AG, Chamerstrasse 56, 6300 Zug, Switzerland; franz.seefried@qualitasag.ch
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-316842529

Simple Summary: Today's Swiss Simmental represents a local dual-purpose breed of cattle. Within closed populations, deleterious variants can reach problematic frequencies, explaining substantial proportions of inbreeding depression. Depletions in homozygous genotypes for certain haplotypes among large cohorts of animals genotyped for the purpose of genomic selection is a widely used approach to pinpoint undesired recessive alleles. In the course of a reverse genetic screen, we aimed to identify single recessive Mendelian variants that potentially affect fertility and rearing success without any phenotypic information available. We detected eleven genome regions showing obvious depletion of homozygosity based on genome-wide SNP data. Furthermore, after performing whole-genome sequencing of selected animals, we propose three candidate causative variants affecting different genes with possibly detrimental effects for embryonic development. The established haplotypes, as well as the identified protein-changing variants, can be directly implemented into breeding practice to avoid the risk of mating carriers and thereby increase breeding success.

Abstract: We herein report the result of a large-scale reverse genetic screen in the Swiss Simmental population, a local dual-purpose cattle breed. We aimed to detect possible recessively inherited variants affecting protein-coding genes, as such deleterious variants can impair fertility and rearing success significantly. We used 115,000 phased SNP data of almost 10 thousand cattle with pedigree data. This revealed evidence for 11 genomic regions of 1.17 Mb on average, with haplotypes (SH1 to SH11) showing a significant depletion in homozygosity and an allele frequency between 3.2 and 10.6%. For the proposed haplotypes, it was unfortunately not possible to evaluate associations with fertility traits as no corresponding data were available. For each haplotype region, possible candidate genes were listed based on their known function in development and disease. Subsequent mining of single-nucleotide variants and short indels in the genomes of 23 sequenced haplotype carriers allowed us to identify three perfectly linked candidate causative protein-changing variants: a SH5-related *DIS3*:p.Ile678fs loss-of-function variant, a SH8-related *CYP2B6*:p.Ile313Asn missense variant, and a SH9-related *NUBPL*:p.Ser143Tyr missense variant. None of these variants occurred in homozygous state in any of more than 5200 sequenced cattle of various breeds. Selection against these alleles in order to reduce reproductive failure and animal loss is recommended.

Keywords: cattle; *Bos taurus*; reproduction; breeding; fertility; embryonic lethality; loss-of-function variants; whole-genome sequencing; SNP genotyping

Citation: Häfliger, I.M.; Seefried, F.R.; Drögemüller, C. Reverse Genetic Screen for Deleterious Recessive Variants in the Local Simmental Cattle Population of Switzerland. *Animals* **2021**, *11*, 3535. https://doi.org/10.3390/ani11123535

Academic Editor: Michael E. Davis

Received: 4 November 2021
Accepted: 8 December 2021
Published: 12 December 2021

1. Introduction

Simmental is a globally recognized cattle breed, originating in the Simmental valley in the canton Bern in Switzerland. Autochthonous of Switzerland, today up to 40 million cattle worldwide are designated as members of the Simmental breed. However, Simmental are characterized by very different local breeding objectives (http://wsff.info; accessed on 24 October 2021). In the 19th century, crossbreeding of local cattle with Simmental cattle exported from Switzerland led to Fleckvieh populations in neighboring

countries, including Austrian and German Fleckvieh, French Montbeliarde, and Italian Pezzata Rossa cattle. In the 20th century, introduction of animals of the Holstein breed into the Swiss Simmental population led to today's Fleckvieh population in Switzerland (Swiss Fleckvieh). A recent analysis of the population structure of Swiss cattle showed a clear differentiation between today's Swiss Simmental cattle and all other Swiss cattle populations, including the modern Swiss Fleckvieh [1,2]. Various historically younger Central European Simmental populations that descended from the Swiss Simmental can also be clearly distinguished from today's Swiss Simmental animals, which showed the highest inbreeding level [3]. In contrast, recent studies found a comparatively low degree of genomic inbreeding in purebred Swiss Simmental that might be explained by the continued use of natural service sires, which is likely the major reason for their remarkably high level of genetic diversity although this population has been closed for a long time [1,2]. Therefore, the current so-called Original Simmental breed of Switzerland represents a unique purebred population. At the end of 2020, more than 23 thousand dual-purpose animals in the Simmental population were registered in the Swiss herdbook (https://www.swissherdbook.ch/fileadmin/Domain1/PDF_Dokumente/05-Statistiken-Formulare/53-Jahresstatistik/1_Wichtigste_Zahlen/D_sh_JS_2020_HBZahlen_web.pdf; accessed on 24 October 2021).

A universal problem in cattle breeding is reproductive failure. It was shown that the reproduction success is negatively associated with production traits [4,5]. These effects have been thoroughly studied and possible reasons are selection programs focusing on production traits coupled with negative correlation with reproductive traits [6]. Regarding early pregnancy loss in Simmental cattle, e.g., a Croatian study described the body condition score, parity, and milk yield as important influencing factors [7]. Lower reproductive success is an economic problem for farmers and a major reason for cattle slaughter [8,9]. To tackle fertility issues, one approach is to apply a so-called reverse genetic screen, where only genomic data is used rather than phenotypic information. A whole-genome sequencing (WGS)-based approach was suggested by Charlier et al. (2016), where sequencing data of entire genomes is mined for variants affecting fertility by causing embryonic lethality [10]. A similar approach proposes scanning the cumulative population-wide SNP genotyping data from genomic selection programs to identify haplotype regions indicative of a depletion in homozygosity [11]. The analysis includes the statistical evaluation of segregating haplotypes regarding the Hardy–Weinberg equilibrium (HWE) [11]. Such genomic regions harboring deficient homozygous haplotypes indicate potential causal variants for embryonic death or unwanted phenotypes in newborns, leading to the exclusion of homozygous animals from breeding programs [11]. This SNP-based reverse genetics approach has been used successfully in diverse breeds and species (e.g., [11–31]). Haplotypes never observed in homozygous state indicate the presence of recessive, predominantly embryonic lethal variants. Potential causative variants were detected by analyzing whole-genome or whole-exome sequencing data linked to the identified haplotypes (e.g., [10,19–30,32–35]). In general, these approaches are especially interesting for small populations where phenotypic information is sparse.

In Simmental-derived breeds in Central Europe, eight causal protein-changing variants for recessively inherited disorders are known so far. In German Fleckvieh cattle, a SNP-based reverse genetic study identified four deficient homozygous haplotypes located on chromosomes 1, 10, and 12 and proposed two candidate causal variants in two genes: the FH2-related frameshift variant in *SLC2A2* and the FH4-related missense variant in *SUGT1* [21]. In addition, a fifth haplotype is described, affecting the calf survival due to a congenital heart failure and severe liver damage (https://www.lfl.bayern.de/itz/rind/122227/index.php; accessed on 24 October 2021). Based on forward genetic studies, where affected individuals were examined, several causal variants for Mendelian disorders were found. A frameshift variant in *GON4L* was associated with the autosomal recessively inherited disorder dwarfism (OMIA 001985-9913) [36]. Furthermore, a nonsense variant in *PLD4* causing a recessive genodermatosis observed in German Fleckvieh (OMIA 001935-

9913) and a missense variant in *OPA3* associated with a form of dilated cardiomyopathy predominantly affecting Swiss Fleckvieh (OMIA 000162-9913) were described [37,38]. In German Fleckvieh, a *MOCS1*-related form of arachnomelia was identified to be due to a frameshift variant (OMIA 001541-9913) [39]. Furthermore, the BH2-related *TUBD1* missense variant, known to cause juvenile mortality in Braunvieh cattle (OMIA 001939-9913), was also observed in German and Austrian Fleckvieh [23]. To the best of our knowledge, so far, no carriers for any of these eight deleterious alleles have been found in the purebred Swiss Simmental population.

In purebred Swiss Simmental cattle, no genomic analysis has yet been carried out to systematically identify recessively inherited harmful variants that affect female reproduction or calf rearing. Based on a reverse genetic screen, the present study aimed at a comprehensive analysis of available SNP and WGS data to identify genomic regions containing deficient homozygous haplotypes as well as linked candidate causal variants for hidden phenotypes.

2. Materials and Methods

The national breeding association provided us with the available SNP data, including all genotyped purebred Swiss Simmental cattle born after 2009 and their ancestors. The genomic positions of the markers relate to the latest cattle reference sequence ARS-UCD1.2 [40,41]. Due to the application of several routinely available SNP arrays ranging from 9000 to 150,000 SNPs, the data had to be imputed. Therefore, the software Fimpute v2.2 [42] was used with default parameters to increase the number of markers and correct for wrongly called markers. In order to assure the quality, SNPs with a minor allele frequency <0.01 were excluded from the dataset. Furthermore, SNPs that could not pass the quality measures for the call rate per SNP >0.99 were excluded and animals with call rates <0.8 were excluded too. Quality control was applied before and after imputation. This resulted in a final SNP data set of 114,890 markers for 9965 animals.

As a first step, haplotypes showing a deviation from the HWE, indicated by depletion of homozygosity, were identified. The first subset of data analyzed included only fully genotyped trios where the complete trio (sire, dam, and offspring) were genotyped (n = 2626), further called "trio" approach. The second dataset analyzed included genotyped trios where an offspring and two paternal animals (sire and maternal grandfather) were genotyped (n = 3969), subsequently called parent–grandparent "pgp" approach. Both data sets, trio and pgp, were used in screening for window size of 50 markers within the software snp1101 [43]. The analyzed haplotypes overlap, as the windows were continuously moved marker by marker. The snp1101 software used the Fisher exact test of HWE [44] to analyze the resulting haplotypes. Furthermore, the p-values were corrected for a false discovery rate with the Benjamini–Yekultieli method and a significance level of 5% was applied [45]. For each significant haplotype region, we selected the most significant haplotype (lowest p-value). Regarding the identification of candidate variants, haplotype regions were defined by the selected haplotypes and increased on both sides by 2 Mb in order to make sure that the regions of overlapping significant haplotypes were included.

Within the 11 regions detected, between 14 and 262 genes were annotated, giving a total of 669 genes. To these genes, information regarding associated phenotypes were collected from the various online databases: HUGO Gene Nomenclature Committee (https://www.genenames.org; accessed on 7 October 2021), Mouse Genome Informatics (http://www.informatics.jax.org; accessed on 7 October 2021), International Mouse Phenotyping Consortium (https://www.mousephenotype.org; accessed on 7 October 2021), Online Mendelian Inheritance in Man (https://www.omim.org; accessed on 7 October 2021), Online Mendelian Inheritance in Animals (https://omia.org/home/; accessed on 7 October 2021), and Decipher (https://www.deciphergenomics.org; accessed on 7 October 2021).

With the selected haplotypes, we predicted individual diplotypes that represent if an animal carries one, two, or no copies of the haplotype. Based on these diplotypes, we selected three carrier animals for whole-genome sequencing (WGS) for each haplotype

region. Therefore, 23 Simmental cattle (1 female and 22 male) carrying one or more significant haplotypes were selected for this study. WGS data were prepared as previously described [46]. However, recalibration was performed with the variant catalogue of the 1000 Bull Genomes Project run 7 (BQSR file version 2) [47,48]. The 23 genomes sequenced in this study were submitted to the 1000 Bull Genomes project [47,48] and are therefore part of the 5116 animals in the recent variant catalogue (run9). This international dataset was used to evaluate candidate causative variants in a larger cohort and across breeds, to evaluate breed specificity as well as reduced homozygosity. Furthermore, we had access to an additional 115 publicly available genomes from the Swiss Comparative Bovine Resequencing project, deposited in the European Nucleotide Archive under project accession PRJEB18113, that were yet not added to the 1000 Bull Genomes project. The combined WGS dataset of 5231 genomes includes 62 purebred Swiss Simmental cattle.

In order to identify haplotype-associated candidate causal variants, linkage disequilibrium (LD) analysis was performed using plink v1.9 [49]. LD (r^2) was calculated between the diplotype state and the WGS-derived genotypes for the 23 animals that were selected for WGS. Analysis was restricted to protein-changing variants, including variants annotated by SnpEff (version 4.3; [50]) with the following sequence ontology terms: missense_variant, start_lost, stop_gained (nonsense), stop_lost, stop_retained_variant, splice_acceptor_variant, splice_donor_variant, conservative_inframe_deletion, conservative_inframe_insertion, disruptive_inframe_deletion, disruptive_inframe_insertion, exon_loss_variant, frameshift_variant, and gene_fusion.

Furthermore, to improve the understanding of candidate causal variants, their base conservation scores from UCSC database called PhyloP and PhastCons were applied [51,52]. To apply these values, firstly, the variants of the bovine genome needed to be mapped to the human genome 38 [53] with the tool LiftOver of the UCSC tools. Secondly, the conservation scores of 99 vertebrates for these human positions were obtained. Additionally, the effects of protein-changing variant were estimated using PROVEAN [54] and PredictSNP [55].

3. Results

In this reverse genetic study, we used SNP data to identify haplotypes showing a depletion in homozygosity and applied WGS data to pinpoint candidate causal variants in purebred Swiss Simmental animals. Reduced homozygosity due to hidden recessive variants in cattle could, so far, only be causally explained by coding variants. Therefore, we focused on variants having a moderate impact, such as missense variants, conservative inframe insertions and deletions up to a size of 50 bp, as well as on all other protein-changing with high impact including loss-of-function variants, such as stop-gains (nonsense), splice site-disrupting SNVs, frameshift indels in a coding sequence, or deletions that remove coding exons.

3.1. Identification of Deficient Homozygous Haplotypes in Swiss Simmental Cattle

We detected seven haplotype regions with the trio approach and nine haplotype regions applying the pgp approach (Table 1, Figure 1, Table S1). Five of these haplotype regions appear in both analyses. We named the haplotype regions, in accordance to previous studies performed in other cattle breeds, as Simmental Haplotypes (SH) 1 to 11 [22] (Table 1, Table S1). All 11 identified haplotypes show a deficiency of at least 85 percent of the expected homozygous animals within the studied Swiss Simmental population (Table 1, Table S1). Four selected haplotypes (SH5, SH7, SH8, and SH10) presented a complete deficit of observed homozygous animals, whereas the others showed a partial deficiency ranging from 85 to 96% of the expected homozygotes (Table 1; Table S1). The average length of the eleven haplotypes is 1.17 Mb and ranges from 0.73 to 1.94 Mb (Table 1).

Table 1. List of deficient homozygous haplotypes in Swiss Simmental cattle.

Haplotype [a]	Analysis	Chr	Position (Mb) [b]	Length (Mb)	Number of Homozygotes			
					Observed	Expected	Deficiency (%)	Allele Frequency
SH1	trio and pgp	1	65.793–66.891	1.10	3	17	85	0.042
SH2	trio	2	1.191–1.920	0.73	1	16	94	0.040
SH3	pgp	7	104.784–105.640	0.86	1	17	94	0.041
SH4	pgp	11	29.389–30.353	0.96	4	22	85	0.047
SH5	trio	12	46.828–47.837	1.01	0	38	100	0.062
	pgp		46.831–47.843	1.01				
SH6	pgp	14	11.304–12.373	1.07	2	15	88	0.039
SH7	trio	16	6.120–8.011	1.89	0	30	100	0.055
	pgp		6.122–8.060	1.94	5	111	96	0.106
SH8	pgp	18	48.763–50.005	1.24	0	11	100	0.0337
	trio		48.806–50.017	1.21				0.0338
SH9	trio and pgp	21	41.855–42.851	1.00	2	19	96	0.043
SH10	pgp	24	41.945–43.140	1.20	0	10	100	0.032
SH11	trio	28	38.937–40.069	1.13	1	12	92	0.035

[a] SH meaning Simmental Haplotype. [b] in accordance with the reference sequence ARS-UCD1.2.

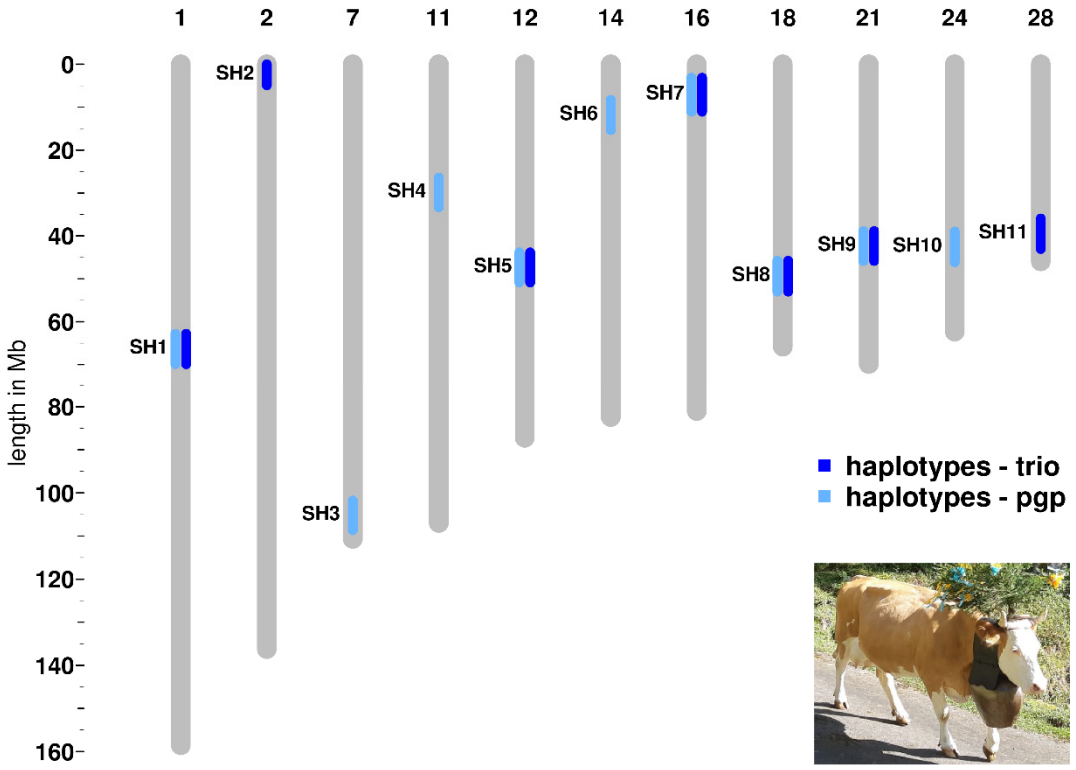


Figure 1. Genomic co-localization of the deficient homozygous haplotype regions detected in Swiss Simmental cattle using two different approaches. A typical breed-specific cow is shown in the lower right corner.

3.2. Identification of Candidate Genes in Haplotype Regions

The intensive analysis of all annotated protein-coding genes in the defined haplotype regions, extended by 2 Mb on each side, led to a comprehensive list of candidate genes possibly affecting either prenatal or postnatal lethality or associated sub-lethal phenotypes. We extracted 145 positional candidate genes of special interest, as they are associated with mammalian autosomal recessive disorders in human, mice, or other animals and listed in the consulted databases (Table S2). Loss-of-function mouse models of many of these

genes have revealed defects that affect embryonic or perinatal to pre-weaning survival and therefore represent suitable functional candidates for this study. The presented short list of the 43 most probable candidate genes includes all genes that are associated with sub-lethal or lethal phenotypes (Table 2).

Table 2. List of candidate genes known to cause recessive disorders with lethal and sub-lethal phenotypes in human.

Haplotype	Gene	Gene Description	Recessive Disorder
SH1	*IQCB1* §	IQ motif containing B1	Senior-Loken syndrome 5
	MYLK §	myosin light chain kinase	Megacystis-microcolon-intestinal hypoperistalsis syndrome 1
	CASR *	calcium sensing receptor	Hyperparathyroidism, neonatal
	POGLUT1 *	protein O-glucosyltransferase 1	Limb-girdle muscular dystrophy-dystroglycanopathy
	TIMMDC1 *	translocase of inner mitochondrial membrane domain containing 1	Mitochondrial complex I deficiency, nuclear type 31
SH2	*HERC2* *,§	HECT and RLD domain containing E3 ubiquitin protein ligase 2	Mental retardation, MRT 38
	OCA2 *,§	OCA2 melanosomal transmembrane protein	Albinism, brown oculocutaneous
SH4	*CRIPT* *	CXXC repeat containing interactor of PDZ3 domain	Short stature with microcephaly and distinctive facies
	EPCAM *	epithelial cell adhesion molecule	Congenital diarrhea with tufting enteropathy
	FSHR *,§	follicle stimulating hormone receptor	Ovarian dysgenesis 1
	MSH2 §	mutS homolog 2	Mismatch repair cancer syndrome 2
	MSH6	mutS homolog 6	Mismatch repair cancer syndrome 3
	PIGF *	phosphatidylinositol glycan anchor biosynthesis class F	Onychodystrophy, osteodystrophy, impaired intellectual development, and seizures syndrome
	PPP1R21 *	protein phosphatase 1 regulatory subunit 21	Neurodevelopmental disorder with hypotonia, facial dysmorphism, and brain abnormalities
	TTC7A *	tetratricopeptide repeat domain 7A	Gastrointestinal defects and immunodeficiency syndrome
SH5	*PIBF1* *	progesterone immunomodulatory binding factor 1	Joubert syndrome 33
SH6	*MYC* *	MYC proto-oncogene, bHLH transcription factor	Burkitt lymphoma, somatic
SH7	*CFH*	complement factor H	Basal laminar drusen, complement factor H deficiency
	CR2 §	complement C3d receptor 2	Immunodeficiency, CVID7
	IL10 §	interleukin 10	Critical role in the control of immune responses
SH8	*ARHGEF1* §	Rho guanine nucleotide exchange factor 1	Immunodeficiency 62
	BCKDHA *	branched chain keto acid dehydrogenase E1 subunit alpha	Maple syrup urine disease, type Ia @
	B9D2 *	B9 domain containing 2	Meckel syndrome 10, Joubert syndrome 34
	COQ8B *	coenzyme Q8B	Nephrotic syndrome, type 9
	DLL3 §	delta-like canonical Notch ligand 3	Spondylocostal dysostosis 1
	ERF *	ETS2 repressor factor	Spondylocostal dysostosis 1
	ETHE1 §	ETHE1 persulfide dioxygenase	Ethylmalonic encephalopathy

Table 2. *Cont.*

Haplotype	Gene	Gene Description	Recessive Disorder
	LTBP4 §	latent transforming growth factor beta binding protein 4	Cutis laxa, autosomal recessive, type IC
	MEGF8 *	multiple EGF-like domains 8	Carpenter syndrome 2
	PLEKHG2	pleckstrin homology and RhoGEF domain containing G2	Leukodystrophy and acquired microcephaly
	RYR1 *	ryanodine receptor 1	Neuromuscular disease
	SMG9 *	SMG9 nonsense mediated mRNA decay factor	Heart and brain malformation syndrome
	SPINT2 *	serine peptidase inhibitor, Kunitz type 2	Diarrhea 3, secretory sodium, congenital, syndromic
	SPTBN4	spectrin beta, non-erythrocytic 4	Neurodevelopmental disorder with hypotonia, neuropathy, and deafness
	TGFB1 *	transforming growth factor beta 1	Inflammatory bowel disease, immunodeficiency, and encephalopathy
	TIMM50 *	translocase of inner mitochondrial membrane 50	3-methylglutaconic aciduria, type IX
	XRCC1 *	X-ray repair cross complementing 1	Spinocerebellar ataxia, SCAR26
SH9	*NUBPL* *	nucleotide binding protein-like	Mitochondrial complex I deficiency, nuclear type 21
SH10	*AFG3L2* *	AFG3-like matrix AAA peptidase subunit 2	Spastic ataxia 5
	LAMA1 *	laminin subunit alpha 1	Poretti–Boltshauser syndrome
	MC2R §	melanocortin 2 receptor	Glucocorticoid deficiency due to ACTH unresponsiveness
	NDUFV2 *	NADH:ubiquinone oxidoreductase core subunit V2	Mitochondrial complex I deficiency, nuclear type 7
	PIEZO2 *	piezo type mechanosensitive ion channel component 2	Arthrogryposis, distal, with impaired proprioception and touch

* Homozygous lethal with complete penetrance, § homozygous lethal with incomplete penetrance, @ disorder described in cattle (OMIA-ID: 000627-9913).

3.3. Identification of Candidate Causal Variants

For three deficient homozygous haplotypes (SH5, SH8, and SH9), by linkage disequilibrium analysis, we found perfectly linked ($r^2 = 1$) candidate causal variants. These three haplotypes were detected with both the pgp and the trio approach. For each of these haplotypes, we propose a protein changing SNV (Table 3; Table S3). These three variants never occur in homozygous state in the analyzed 5231 bovine genomes of various cattle breeds (Table S3). Interestingly, the SH8-associated variant is apparently specific for Swiss Simmental; however, the variants associated with SH5 and SH9 occur sporadically in single animals of some other breeds (Table S3).

Table 3. Candidate causal variants for deficiency of homozygotes in Swiss Simmental cattle.

Haplotype	Gene	OMIM	Variant	Transcript [a]	Coding DNA Change	Predicted Protein Change
SH5	*DIS3*	607533	1 bp insertion (frameshift)	XM_025000110.1	c.2032dupA	p.Ile678AsnTer2
SH8	*CYP2B6*	123930	SNV (missense)	NM_001075173.1	c.938T > A	p.Ile313Asn
SH9	*NUBPL*	613621	SNV (missense)	NM_001193042.1	c.428C > A	p.Ser143Tyr

[a] According to the NCBI Annotation Release 106 (National Center for Biotechnology Information, 2018b).

Among the three proposed non-synonymous variants are two missense variants altering evolutionary conserved residues and a frameshift variant that significantly truncates the encoded protein. The SH8-related SNV in exon 4 of the bovine *cytochrome P450 family 2 subfamily B member 6* (*CYP2B6*) gene on chromosome 18 at position 50296371 is a missense variant (NM_001075173.1: p.Ile313Asn) that was predicted by PROVEAN to have a deleterious effect (Table 3, Table S3). The SH9-related SNV located in exon 6 of the bovine *nucleotide binding protein-like* (*NUBPL*) gene on chromosome 21 at position 42154344 represents a missense variant (NM_001193042.1: p.Ser143Tyr) predicted to be deleterious by PROVEAN and PredictSNP (Figure 2, Table 3, Table S3). Both presented missense mutations altering evolutionary conserved amino acids (Figure 2). The SH5-associated 1 bp insertion located in exon 16 of the bovine *DIS3 homolog, exosome endoribonuclease and 3′-5′ exoribonuclease* (*DIS3*) gene on chromosome 12 at position 47511687 represents a loss-of-function variant. It was predicted to result in a frameshift after isoleucine 678 with a premature stop codon (NP_025000110.1: p.Ile678AsnTer2), resulting in a significantly shortened amino acid sequence, if expressed, when compared with the wild-type protein (Table 3; Table S3). In addition, the comparative DNA sequence approach (PhyloP and phastCons) showed a high conservation across species for all three variant positions (Table S3).

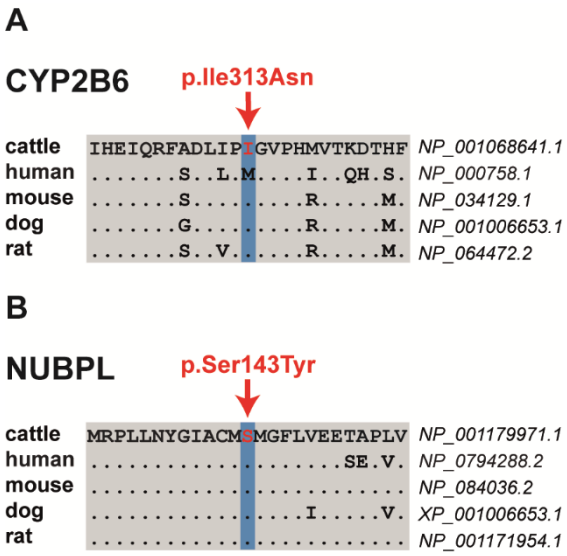


Figure 2. Multiple sequence alignment of the amino acid sequences of CYP2B6 (**A**) and NUBPL (**B**) around the position of the missense variants (shown in red), indicating evolutionary conservation across all species.

4. Discussion

For the first time, the genomic data of the current Swiss Simmental dual-purpose cattle population were analyzed for reduced homozygosity due to hidden recessive monogenic variants and validated using the international variant catalogue of the 1000 Bull Genomes Project. In addition to environmental factors, inherited deleterious variants lead to natural or artificial selection against homozygous individuals, which also explains embryonic lethality, reduce rearing success, or the exclusion from the breeding population due to poor development. Unfortunately, these phenomena are not systematically monitored and are therefore difficult phenotypes to study. Although reduced reproductive success can theoretically be detected in the sires estimated breeding values for certain fertility traits, these effects are only noticeable for deleterious alleles that have reached a high frequency in the population [56]. To overcome this issue, we performed a genome-wide missing homozygosity scan, revealing eleven haplotype regions with considerable homozygous depletion. After subsequent mining of genome sequence data for candidate causal variants, we propose three non-synonymous variants that probably cause the obvious deficiency of homozygous animals.

We applied two different approaches, a trio and pgp approach, that include genotyped trios and applies a Fisher exact test. Therefore, we were able to detect haplotypes that segregate at a lower allele frequency. As expected, we found more haplotype regions with the pgp approach in comparison to the trio approach, most likely because more genotyped groups were available. Nevertheless, the trio approach, detecting two regions not found with pgp, appears very powerful, probably because it directly traces the inheritance of the haplotypes. Previous approaches used the assumption of random mating and the deviation from HWE based on allele frequencies or used the deviation from expected number of homozygous offspring based on the haplotype state of the sire and maternal grandsire [14,20]. In contrast, the herein applied trio approach allows performing such an analysis with a small population such as Simmental. The five haplotype regions that were found in both analyses are the most probable genome regions harboring hidden harmful variants. Especially the haplotypes that never occur in homozygous state led to the suspicion of embryonic lethal variants segregating in the population. It is suspected that some haplotypes arose due to imputation errors introduced due to genotyping bias, SNP density/panel, sample size, and a bias introduced by the chosen software [57]. This would explain the haplotype regions SH3 and SH11 that arose in regions for which we could not identify any plausible functional candidate genes. Otherwise, we were able to identify candidate genes within all missing homozygous regions. Unfortunately, effect estimations of the haplotypes towards traits of female fertility and rearing success were not reasonable, as the available phenotypic data is currently limited. To avoid the detection of mostly sporadic associations rather than actual effects, it is planned to conduct such haplotype association studies in the future.

Candidate causal variants are proposed for the haplotypes SH5, SH8, and SH9 in the genes *DIS3*, *CYP2B6,* and *NUBPL*, respectively. These variants all show complete depletion in homozygosity, perfect LD to the associated haplotype, and high conservation scores when compared across 99 genomes, indicating their importance in basic biological functions. As in other mammals, it is expected that ~100 harmful recessive variants will be found per individual in cattle, of which up to five of these impact essential genes and cause embryonic lethality or severe disease when homozygous [10]. Nevertheless, it is recognized that it is very difficult to clarify the actual deleterious functions of these variants, although given the genes involved, it is assumed that these variants influence fitness. Reverse genetic screens to identify genes with major effects, as used in the current study, are therefore helpful to assign function to variants in candidate genes and/or so far less characterized genes such as *DIS3* and *CYP2B6*.

The SH5-related loss-of-function variant found in bovine *DIS3,* most likely leading to missing homozygosity, represents the first time a pathogenic variant that most likely causes embryonic lethality has been identified. If the mutant mRNA transcript were to escape nonsense-mediated decay, even if this truncated protein was expressed, it would lack roughly 30 percent of the C-terminal part, and therefore it is not expected to contribute any function. *DIS3*, also known as ribosomal RNA-processing protein 44 (RRP44), is a RNase II/R-like enzyme located primarily in the nucleus (Table S3) [58]. The protein has catalytic function in the RNA exosome complex, which is responsible for 3′-end processing and RNA degradation of a broad variety of RNAs [58,59]. Biological functions are associated with RNA metabolism, mitotic control, spindle-fiber formation, antibody diversification, microtubule production, and growth and development [60]. Recently, pathogenic variants in genes encoding both structural and catalytic subunits of the RNA exosome have been linked to human disease, such as *EXOSC3* and *EXOSC8* related forms of pontocerebellar hypoplasia, representing recessive neurodegenerative diseases [61]. To our knowledge, the *DIS3* gene has not yet been associated with Mendelian diseases, but variants are reported to be associated with various types of cancers and multiple myeloma [60–62]. Interestingly, the book of Fasken et al. (2020) provides a summary of the most common variants in *DIS3* which all occur in heterozygous state, are associated with multiple myeloma, and seem to have mild effects only, while Tomecki et al. (2014) suggests the potential lethality of

mutations in the PIN domain of *DIS3*. In *Drosophila*, a knock-down model led to wingless animals implicating an important role in development [63]. Nevertheless, the inactivation of *DIS3* in B cells was shown to lead to an increase in unbalanced DNA translocations [64]. Lastly, public databases for mouse phenotypes indicate a complete pre-weaning lethality of *DIS3* knock-out mice (https://www.mousephenotype.org/data/genes/MGI:1919912; http://www.informatics.jax.org/marker/MGI:1919912; accessed on 7 October 2021).

Despite a long list of candidate genes located within the SH8-region, we propose a missense variant in the *CYP2B6* gene as a candidate causal variant. The main reason for this is the perfect LD to the haplotype, the complete absence of homozygous animals, and the prediction of the DNA position to be highly conserved and the amino acid exchange to be deleterious. Regarding the gene function, CYP2B6 is a protein of the cytochromes P450 subfamily 2B (HGNC: 20604). This enzyme is known to be of importance for drug metabolism, as well as endogenous compounds, environmental toxins, and other substances [65,66]. For example, the susceptibility to Efavirenz depends on the individual *CYP2B6* genotype (OMIM: 123930). Several SNV were detected in human that are associated with the expression level and activity (increased and decreased) of *CYP2B6* with a population-wide importance [65] (https://www.pharmvar.org/gene/CYP2B6; accessed on 26 October 2021). In monkeys and human, it was shown that *CYP2B6* is expressed in the brain and affected by nicotine and alcohol consumption; however, its neurological function remains unclear [65–67]. The protein is also expressed in the placenta [68] and it was shown that the pregnancy hormone estradiol induces the expression of *CYP2B6* [69]. Nevertheless, what the function and importance of *CYP2B6* is in maintaining pregnancy is unclear. As the herein identified bovine variant is predicted to be deleterious, we speculate that function of CYP2B6 might be impaired during development.

Lastly, we propose the missense variant in the bovine *NUBPL* gene, exchanging a strongly conserved residue predicted to be deleterious and affecting a highly conserved nucleotide, to be causal for the deficit of homozygosity of SH9. As we observe few haplotype carriers in our data, we speculate that the effect of the variant is not fully penetrant or that signs of poor development appear later in life, after initial genotyping of young animals. NUBPL, also known as *iron-sulfur protein required for NADH dehydrogenase* (*IND1*), is a protein that is vital for the assembly of the respiratory complex I [70]. More precisely, the NUBPL supplies Fe/S clusters to the respiratory complex I and thereby ensures that important subunits are delivered to build the whole complex [70]. In human, pathogenic variants affecting *NUBPL* are associated with the autosomal recessive mitochondrial complex I deficiency disorder (OMIM: 613621, 618242) [71–73]. These variants include missense, frameshift and splice site variants, as well as small and large insertions and deletions. Clinical symptoms of mitochondrial complex I deficiency include, among others, ataxia, dysarthria, hypotonia, nystagmus, spasticity, and tremor [73]. Furthermore, variants in *NUBPL* were hypothesized as risk factors for Parkinson's disease [74]. A mouse model identified the necessity of the protein as knock-out alleles led to homozygous lethality (MGI: 1924076) [75].

However, if the proposed variants are indeed depleted in a number of homozygous animals, it needs to be evaluated by further genotyping of larger cohorts before implementation into selection schemes. This is planned by adding the variants to a custom array for genotyping larger cohorts of further animals, both Swiss Simmental as well as other local cattle populations. An alternative approach to confirm the absence of homozygous animals, particularly in the offspring of carrier-by-carrier mating's, would strongly support the deleterious nature of the variants as shown before [10]. In particular, for the *NUBPL*-associated variant, the observed homozygous animals should be examined in detail.

The chosen approach requires that the haplotype and the causative variant are in near perfect linkage disequilibrium and, obviously, this is not always the case. This could be an explanation for the fact that we could not identify any potentially causal variants in the other haplotype regions. An alternative approach, therefore, is to mine the genome sequence data for candidate variants, e.g., loss-of-function in the listed essential candidate

genes, and to genotype these directly in large cohorts of Swiss Simmental cattle. Using this approach, nine causal variants were uncovered in cattle that would not have been detected using SNP-based haplotype approaches [10].

Finally, another drawback of our study was the restriction to protein-changing variants. Moreover, we had no evidence for perfectly linked non-coding regulatory variants, as well as the limitation to consider only SNV and short indels, overlooking possible larger structural variants.

5. Conclusions

In the presented project, we mined SNP and WGS data by applying a reverse genetic approach. Without any phenotypic evidence but mining the data of almost 10 thousand SNP genotyped Swiss Simmental cattle and more than 5200 WGS animals from a variety of breeds, we propose three candidate variants in the genes *DIS3*, *CYP2B6,* and *NUBPL* causing embryonic lethality and/or yet unknown recessive developmental disorders. After phenotypic validation of these variants, selection against these variants is recommended.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ani11123535/s1, Table S1: Haplotype information of SH1 to SH11, Table S2: Comprehensive list of candidate genes located in the haplotype regions, including gene information from MGI, IMPC, OMIM, and OMIA, Table S3: Comprehensive list of potential candidate causal variants.

Author Contributions: Conceptualization, C.D. and F.R.S.; formal analysis, I.M.H. and F.R.S.; investigation, I.M.H. and F.R.S.; data curation, I.M.H.; writing—original draft preparation, I.M.H.; writing—review and editing, C.D., F.R.S.; visualization, I.M.H.; supervision, C.D. and F.R.S.; project administration and funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Swiss National Science Foundation, grant number 172911.

Institutional Review Board Statement: This study did not require official or institutional ethical approval as it was not an experimental study, but an analysis of existing genotype data.

Data Availability Statement: The SNP data of Swiss Simmental cattle are owned by the breeding association swissherdbook. Therefore, we ask interested people to contact the authors or the breeding association directly in order to gain access to the SNP data. The WGS data are publicly available at the European Nucleotide Archive under project accession PRJEB18113.

Acknowledgments: We are very thankful for the genomic data provided by the Swiss cattle breeding association swissherdbook. The authors also wish to thank Swissgenetics and Selectstar for providing semen. The authors are thankful to the University of Bern institutions, namely the Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit, for performing the WGS experiments and providing high performance computing infrastructure, respectively. Finally, we acknowledge the 1000 Bull Genomes Project consortium, led by Christy Vander Jagt, Amanda Chemberlain, and Hans Daetwyler, for providing the variant catalogue.

Conflicts of Interest: The authors declare no conflict of interest.

Ethics Approval and Consent to Participate: According to the Swiss laws, no ethics approval was needed. In this study, available data were reused and mined using novel approaches.

References

1. Signer-Hasler, H.; Burren, A.; Neuditschko, M.; Frischknecht, M.; Garrick, D.; Stricker, C.; Gredler, B.; Bapst, B.; Flury, C. Population structure and genomic inbreeding in nine Swiss dairy cattle populations. *Genet. Sel. Evol.* **2017**, *49*, 83. [CrossRef]
2. Stranzinger, G.F.; Steiger, D.; Kneubühler, J.; Hagger, C. Y chromosome polymorphism in various breeds of cattle (*Bos taurus*) in Switzerland. *J. Appl. Genet.* **2007**, *48*, 241–245. [CrossRef]
3. Cesarani, A.; Gaspa, G.; Pauciullo, A.; Degano, L.; Vicario, D.; Macciotta, N.P.P. Genome-wide analysis of homozygosity regions in european simmental bulls. *J. Anim. Breed. Genet.* **2021**, *138*, 69–79. [CrossRef]
4. Pryce, J.E.; Woolaston, R.; Berry, D.P.; Wall, E.; Winters, M.; Butler, R.; Shaffer, M. World Trends in Dairy Cow Fertility. In Proceedings of the 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada, 17–22 August 2014; Volume 10, p. 6.

5. Walsh, S.W.; Williams, E.J.; Evans, A.C.O. A review of the causes of poor fertility in high milk producing dairy cows. *Anim. Reprod. Sci.* **2011**, *123*, 127–138. [CrossRef] [PubMed]
6. Ma, L.; Sonstegard, T.S.; Cole, J.B.; Vantassell, C.P.; Wiggans, G.R.; Crooker, B.A.; Tan, C.; Prakapenka, D.; Liu, G.E.; Da, Y. Genome changes due to artificial selection in U.S. Holstein cattle. *BMC Genom.* **2019**, *20*, 128. [CrossRef] [PubMed]
7. Zobel, R.; Tkalčić, S.; Pipal, I.; Buić, V. Incidence and factors associated with early pregnancy losses in Simmental dairy cows. *Anim. Reprod. Sci.* **2011**, *127*, 121–125. [CrossRef]
8. Seegers, H.; Beaudeau, F.; Fourichon, C.; Bareille, N. Reasons for culling in French Holstein cows. *Prev. Vet. Med.* **1998**, *36*, 257–271. [CrossRef]
9. Rilanto, T.; Reimus, K.; Orro, T.; Emanuelson, U.; Viltrop, A.; Mõtus, K. Culling reasons and risk factors in Estonian dairy cows. *BMC Vet. Res.* **2020**, *16*, 173. [CrossRef]
10. Charlier, C.; Li, W.; Harland, C.; Littlejohn, M.; Coppieters, W.; Creagh, F.; Davis, S.; Druet, T.; Faux, P.; Guillaume, F.; et al. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. *Genome Res.* **2016**, *26*, 1333–1341. [CrossRef] [PubMed]
11. VanRaden, P.M.; Olson, K.M.; Null, D.J.; Hutchison, J.L. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* **2011**, *94*, 6153–6161. [CrossRef]
12. Casellas, J.; Canas-Alvarez, J.J.; Gonzalez-Rodriguez, A.; Puig-Oliveras, A.; Fina, M.; Piedrafita, J.; Molina, A.; Diaz, C.; Baro, J.A.; Varona, L. Bayesian analysis of parent-specific transmission ratio distortion in seven Spanish beef cattle breeds. *Anim. Genet.* **2017**, *48*, 93–96. [CrossRef] [PubMed]
13. Casellas, J.; Gularte, R.J.; Farber, C.R.; Varona, L.; Mehrabian, M.; Schadt, E.E.; Lusis, A.J.; Attie, A.D.; Yandell, B.S.; Medrano, J.F. Genome Scans for Transmission Ratio Distortion Regions in Mice. *Genetics* **2012**, *191*, 247–259. [CrossRef] [PubMed]
14. Derks, M.F.L.; Megens, H.J.; Bosse, M.; Lopes, M.S.; Harlizius, B.; Groenen, M.A.M. A systematic survey to identify lethal recessive variation in highly managed pig populations. *BMC Genom.* **2017**, *18*, 858. [CrossRef]
15. Derks, M.F.L.; Megens, H.J.; Bosse, M.; Visscher, J.; Peeters, K.; Bink, M.; Vereijken, A.; Gross, C.; de Ridder, D.; Reinders, M.J.T.; et al. A survey of functional genomic variation in domesticated chickens. *Genet. Sel. Evol.* **2018**, *50*, 17. [CrossRef]
16. Georges, M.; Charlier, C.; Hayes, B. Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* **2019**, *20*, 135–156. [CrossRef]
17. Abdalla, E.A.; Id-Lahoucine, S.; Cánovas, A.; Casellas, J.; Schenkel, F.S.; Wood, B.J.; Baes, C.F. Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. *Anim. Genet.* **2020**, *51*, 876–889. [CrossRef] [PubMed]
18. Hoff, J.L.; Decker, J.E.; Schnabel, R.D.; Taylor, J.F. Candidate lethal haplotypes and causal mutations in Angus cattle. *BMC Genom.* **2017**, *18*, 799. [CrossRef]
19. Mesbah-Uddin, M.; Hoze, C.; Michot, P.; Barbat, A.; Lefebvre, R.; Boussaha, M.; Sahana, G.; Fritz, S.; Boichard, D.; Capitan, A. A missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. *J. Dairy Sci.* **2019**, *102*, 6340–6356. [CrossRef] [PubMed]
20. Fritz, S.; Capitan, A.; Djari, A.; Rodriguez, S.C.; Barbat, A.; Baur, A.; Grohs, C.; Weiss, B.; Boussaha, M.; Esquerre, D.; et al. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS ONE* **2013**, *8*, e65550. [CrossRef]
21. Pausch, H.; Schwarzenbacher, H.; Burgstaller, J.; Flisikowski, K.; Wurmser, C.; Jansen, S.; Jung, S.; Schnieke, A.; Wittek, T.; Fries, R. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genom.* **2015**, *16*, 312. [CrossRef]
22. Cole, J.B.; VanRaden, P.M.; Null, D.J.; Hutchison, J.L.; Cooper, T.A.; Hubbard, S.M. Haplotype Tests for Recessive Disorders that Affect Fertility and Other Traits. 2021. Available online: https://www.aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html (accessed on 15 April 2021).
23. Schwarzenbacher, H.; Burgstaller, J.; Seefried, F.R.; Wurmser, C.; Hilbe, M.; Jung, S.; Fuerst, C.; Dinhopl, N.; Weissenböck, H.; Fuerst-Waltl, B.; et al. A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. *BMC Genom.* **2016**, *17*, 400. [CrossRef] [PubMed]
24. Sonstegard, T.S.; Cole, J.B.; VanRaden, P.M.; Van Tassell, C.P.; Null, D.J.; Schroeder, S.G.; Bickhart, D.; McClure, M.C. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. *PLoS ONE* **2013**, *8*, e54872. [CrossRef] [PubMed]
25. Charlier, C.; Agerholm, J.S.; Coppieters, W.; Karlskov-Mortensen, P.; Li, W.; de Jong, G.; Fasquelle, C.; Karim, L.; Cirera, S.; Cambisano, N.; et al. A Deletion in the Bovine FANCI Gene Compromises Fertility by Causing Fetal Death and Brachyspina. *PLoS ONE* **2012**, *7*, e43085. [CrossRef]
26. McClure, M.; Kim, E.; Bickhart, D.; Null, D.; Cooper, T.; Cole, J.; Wiggans, G.; Ajmone-Marsan, P.; Colli, L.; Santus, E.; et al. Fine Mapping for Weaver Syndrome in Brown Swiss Cattle and the Identification of 41 Concordant Mutations across NRCAM, PNPLA8 and CTTNBP2. *PLoS ONE* **2013**, *8*, e59251. [CrossRef]
27. Schütz, E.; Wehrhahn, C.; Wanjek, M.; Bortfeld, R.; Wemheuer, W.E.; Beck, J.; Brenig, B. The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. *PLoS ONE* **2016**, *11*, e0154602. [CrossRef]

28. Null, D.J.; Huthcins, J.L.; Bickhart, D.M.; VanRaden, P.M.; Cole, J.B. Discovery of a haplotype affecting fertility in Ayrshire dairy cattle and identification of a putative causal variant. In Proceedings of the American Dairy Science Association Annual Meeting, Pittsburgh, PA, USA, 25–29 June 2017; Volume 100, p. 199.
29. Guarini, A.R.; Sargolzaei, M.; Brito, L.F.; Kroezen, V.; Lourenco, D.A.L.; Baes, C.F.; Miglior, F.; Cole, J.B.; Schenkel, F.S. Estimating the effect of the deleterious recessive haplotypes AH1 and AH2 on reproduction performance of Ayrshire cattle. *J. Dairy Sci.* **2019**, *102*, 5315–5322. [CrossRef]
30. Cooper, T.A.; Wiggans, G.R.; Null, D.J.; Hutchison, J.L.; Cole, J.B. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. *J. Dairy Sci.* **2014**, *97*, 3878–3882. [CrossRef]
31. Ben Braiek, M.; Fabre, S.; Hozé, C.; Astruc, J.-M.; Moreno-Romieux, C. Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. *Genet. Sel. Evol.* **2021**, *53*, 41. [CrossRef]
32. Michot, P.; Fritz, S.; Barbat, A.; Boussaha, M.; Deloche, M.C.; Grohs, C.; Hoze, C.; Le Berre, L.; Le Bourhis, D.; Desnoes, O.; et al. A missense mutation in PFAS (phosphoribosylformylglycinamidine synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbeliarde dairy cattle. *J. Dairy Sci.* **2017**, *100*, 8176–8187. [CrossRef]
33. Fritz, S.; Hoze, C.; Rebours, E.; Barbat, A.; Bizard, M.; Chamberlain, A.; Escouflaire, C.; Vander Jagt, C.; Boussaha, M.; Grohs, C.; et al. An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **2018**, *101*, 6220–6231. [CrossRef]
34. Hozé, C.; Escouflaire, C.; Mesbah-Uddin, M.; Barbat, A.; Boussaha, M.; Deloche, M.C.; Boichard, D.; Fritz, S.; Capitan, A. Short communication: A splice site mutation in CENPU is associated with recessive embryonic lethality in Holstein cattle. *J. Dairy Sci.* **2020**, *103*, 607–612. [CrossRef]
35. Kunz, E.; Rothammer, S.; Pausch, H.; Schwarzenbacher, H.; Seefried, F.R.; Matiasek, K.; Seichter, D.; Russ, I.; Fries, R.; Medugorac, I. Confirmation of a non-synonymous SNP in PNPLA8 as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. *Genet. Sel. Evol.* **2016**, *48*, 21. [CrossRef] [PubMed]
36. Schwarzenbacher, H.; Wurmser, C.; Flisikowski, K.; Misurova, L.; Jung, S.; Langenmayer, M.C.; Schnieke, A.; Knubben-Schweizer, G.; Fries, R.; Pausch, H. A frameshift mutation in GON4L is associated with proportionate dwarfism in Fleckvieh cattle. *Genet. Sel. Evol.* **2016**, *48*, 25. [CrossRef] [PubMed]
37. Jung, S.; Pausch, H.; Langenmayer, M.C.; Schwarzenbacher, H.; Majzoub-Altweck, M.; Gollnick, N.S.; Fries, R. A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. *BMC Genom.* **2014**, *15*, 623. [CrossRef] [PubMed]
38. Owczarek-Lipska, M.; Plattet, P.; Zipperle, L.; Drögemüller, C.; Posthaus, H.; Dolf, G.; Braunschweig, M.H. A nonsense mutation in the optic atrophy 3 gene (OPA3) causes dilated cardiomyopathy in Red Holstein cattle. *Genomics* **2011**, *97*, 51–57. [CrossRef]
39. Seichter, D.; Russ, I.; Förster, M.; Medugorac, I. SNP-based association mapping of Arachnomelia in Fleckvieh cattle. *Anim. Genet.* **2011**, *42*, 544–547. [CrossRef]
40. National Center for Biotechnology Information ARS-UCD1.2 Assembly. 2018. Available online: https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1 (accessed on 31 May 2018).
41. Rosen, B.D.; Bickhart, D.M.; Schnabel, R.D.; Koren, S.; Elsik, C.G.; Tseng, E.; Rowan, T.N.; Low, W.Y.; Zimin, A.; Couldrey, C.; et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* **2020**, *9*, giaa021. [CrossRef]
42. Sargolzaei, M.; Chesnais, J.P.; Schenkel, F.S. A new approach for efficient genotype imputation using information from relatives. *BMC Genom.* **2014**, *15*, 478. [CrossRef]
43. Sargolzaei, M. *SNP1101 User's Guide. Version 1.0*; HiggsGene Solut. Inc.: Guelph, ON, Canada, 2014.
44. Wigginton, J.E.; Cutler, D.J.; Abecasis, G.R. A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Genet.* **2005**, *76*, 887–893. [CrossRef]
45. Benjamini, Y.; Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **2001**, *29*, 1165–1188. [CrossRef]
46. Häfliger, I.M.; Wiedemar, N.; Švara, T.; Starič, J.; Cociancich, V.; Šest, K.; Gombač, M.; Paller, T.; Agerholm, J.S.; Drögemüller, C. Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. *Anim. Genet.* **2020**, *51*, 382–390. [CrossRef]
47. Bovine Research Community 1000 Bull Genomes Project. 2018. Available online: http://www.1000bullgenomes.com/ (accessed on 15 June 2018).
48. Hayes, B.J.; Daetwyler, H.D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annu. Rev. Anim. Biosci.* **2019**, *7*, 89–102. [CrossRef] [PubMed]
49. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
50. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]
51. Pollard, K.S.; Hubisz, M.J.; Rosenbloom, K.R.; Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **2010**, *20*, 110–121. [CrossRef]

52. Siepel, A.; Bejerano, G.; Pedersen, J.S.; Hinrichs, A.S.; Hou, M.; Rosenbloom, K.; Clawson, H.; Spieth, J.; Hillier, L.D.W.; Richards, S.; et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **2005**, *15*, 1034–1050. [CrossRef] [PubMed]
53. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; Fitzhugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [CrossRef]
54. Choi, Y.; Sims, G.E.; Murphy, S.; Miller, J.R.; Chan, A.P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* **2012**, *7*, e46688. [CrossRef]
55. Bendl, J.; Stourac, J.; Salanda, O.; Pavelka, A.; Wieben, E.D.; Zendulka, J.; Brezovsky, J.; Damborsky, J. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput. Biol.* **2014**, *10*, e1003440. [CrossRef] [PubMed]
56. Kadri, N.K.; Sahana, G.; Charlier, C.; Iso-Touru, T.; Guldbrandtsen, B.; Karim, L.; Nielsen, U.S.; Panitz, F.; Aamand, G.P.; Schulman, N.; et al. A 660-Kb Deletion with Antagonistic Effects on Fertility and Milk Production Segregates at High Frequency in Nordic Red Cattle: Additional Evidence for the Common Occurrence of Balancing Selection in Livestock. *PLoS Genet.* **2014**, *10*, e1004049. [CrossRef]
57. Shi, S.; Yuan, N.; Yang, M.; Du, Z.; Wang, J.; Sheng, X.; Wu, J.; Xiao, J. Comprehensive Assessment of Genotype Imputation Performance. *Hum. Hered.* **2018**, *83*, 107–116. [CrossRef]
58. Tomecki, R.; Kristiansen, M.S.; Lykke-Andersen, S.; Chlebowski, A.; Larsen, K.M.; Szczesny, R.J.; Drazkowska, K.; Pastula, A.; Andersen, J.S.; Stepien, P.P.; et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.* **2010**, *29*, 2342–2357. [CrossRef]
59. Schmid, M.; Jensen, T.H. *The Nuclear RNA Exosome and Its Cofactors BT—The Biology of mRNA: Structure and Function*; Oeffinger, M., Zenklusen, D., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 113–132. ISBN 978-3-030-31434-7.
60. Robinson, S.R.; Oliver, A.W.; Chevassut, T.J.; Newbury, S.F. The 3′ to 5′ Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. *Biomolecules* **2015**, *5*, 1515–1539. [CrossRef]
61. Fasken, M.B.; Morton, D.J.; Kuiper, E.G.; Jones, S.K.; Leung, S.W.; Corbett, A.H. *The RNA Exosome and Human Disease BT—The Eukaryotic RNA Exosome: Methods and Protocols*; LaCava, J., Vaňáčová, Š., Eds.; Springer New York: New York, NY, USA, 2020; pp. 3–33. ISBN 978-1-4939-9822-7.
62. Tomecki, R.; Drazkowska, K.; Kucinski, I.; Stodus, K.; Szczesny, R.J.; Gruchota, J.; Owczarek, E.P.; Kalisiak, K.; Dziembowski, A. Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. *Nucleic Acids Res.* **2014**, *42*, 1270–1290. [CrossRef]
63. Towler, B.P.; Jones, C.I.; Viegas, S.C.; Apura, P.; Waldron, J.A.; Smalley, S.K.; Arraiano, C.M.; Newbury, S.F. The 3′-5′ exoribonuclease Dis3 regulates the expression of specific microRNAs in Drosophila wing imaginal discs. *RNA Biol.* **2015**, *12*, 728–741. [CrossRef] [PubMed]
64. Laffleur, B.; Lim, J.; Zhang, W.; Chen, Y.; Pefanis, E.; Bizarro, J.; Batista, C.R.; Wu, L.; Economides, A.N.; Wang, J.; et al. Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nat. Genet.* **2021**, *53*, 230–242. [CrossRef]
65. Turpeinen, M.; Zanger, U.M. Cytochrome P450 2B6: Function, genetics, and clinical relevance. *Drug Metabol. Drug Interact.* **2012**, *27*, 185–197. [CrossRef] [PubMed]
66. Lee, A.M.; Miksys, S.; Palmour, R.; Tyndale, R.F. CYP2B6 is expressed in African Green monkey brain and is induced by chronic nicotine treatment. *Neuropharmacology* **2006**, *50*, 441–450. [CrossRef]
67. Miksys, S.; Lerman, C.; Shields, P.G.; Mash, D.C.; Tyndale, R.F. Smoking, alcoholism and genetic polymorphisms alter CYP2B6 levels in human brain. *Neuropharmacology* **2003**, *45*, 122–132. [CrossRef]
68. Li, L.; Zhang, Q.-Y.; Ding, X. A CYP2B6-humanized mouse model and its potential applications. *Drug Metab. Pharmacokinet.* **2018**, *33*, 2–8. [CrossRef]
69. Koh, K.H.; Jurkovic, S.; Yang, K.; Choi, S.-Y.; Jung, J.W.; Kim, K.P.; Zhang, W.; Jeong, H. Estradiol induces cytochrome P450 2B6 expression at high concentrations: Implication in estrogen-mediated gene regulation in pregnancy. *Biochem. Pharmacol.* **2012**, *84*, 93–103. [CrossRef] [PubMed]
70. Sheftel, A.D.; Stehling, O.; Pierik, A.J.; Netz, D.J.A.; Kerscher, S.; Elsässer, H.-P.; Wittig, I.; Balk, J.; Brandt, U.; Lill, R. Human ind1, an iron-sulfur cluster assembly factor for respiratory complex I. *Mol. Cell. Biol.* **2009**, *29*, 6059–6073. [CrossRef] [PubMed]
71. Calvo, S.E.; Tucker, E.J.; Compton, A.G.; Kirby, D.M.; Crawford, G.; Burtt, N.P.; Rivas, M.; Guiducci, C.; Bruno, D.L.; Goldberger, O.A.; et al. High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nat. Genet.* **2010**, *42*, 851–858. [CrossRef]
72. Kevelam, S.H.; Rodenburg, R.J.; Wolf, N.I.; Ferreira, P.; Lunsing, R.J.; Nijtmans, L.G.; Mitchell, A.; Arroyo, H.A.; Rating, D.; Vanderver, A.; et al. NUBPL mutations in patients with complex I deficiency and a distinct MRI pattern. *Neurology* **2013**, *80*, 1577–1583. [CrossRef] [PubMed]
73. Kimonis, V.; Al Dubaisi, R.; Maclean, A.E.; Hall, K.; Weiss, L.; Stover, A.E.; Schwartz, P.H.; Berg, B.; Cheng, C.; Parikh, S. NUBPL mitochondrial disease: New patients and review of the genetic and clinical spectrum. *J. Med. Genet.* **2021**, *58*, 314–325. [CrossRef]

74. Eis, P.S.; Huang, N.; Langston, J.W.; Hatchwell, E.; Schüle, B. Loss-of-Function NUBPL Mutation May Link Parkinson's Disease to Recessive Complex I Deficiency. *Front. Neurol.* **2020**, *11*, 1332. [CrossRef]
75. Dickinson, M.E.; Flenniken, A.M.; Ji, X.; Teboul, L.; Wong, M.D.; White, J.K.; Meehan, T.F.; Weninger, W.J.; Westerberg, H.; Adissu, H.; et al. High-throughput discovery of novel developmental phenotypes. *Nature* **2016**, *537*, 508–514. [CrossRef]

4 Discussion and perspectives

During the work on this thesis, I investigated the genetic background of both hidden and visible phenotypes by applying forward and reverse genetic approaches. Applying FGA led to the identification of candidate causative variants for seven recessive and ten dominant disorders in various breeds sampled across Europe (Table 5 and Table 6, respectively and Table A7). By using the RGA methods on massive genomic and phenotypic data, 115 haplotype regions and 13 novel candidate causative recessive variants potentially affecting reproductive success in the main Swiss dairy cattle breeds were identified (Figure 11 and Table 7; Table A8 and Table A9).

Table 5: The herein described recessively inherited disorders and their associated genes and types of variants identified through the forward genetic approach.

associated disorder	gene	identified variant	breed	AF (%)	OMIA	ref
Achromatopsia (OH1)	*CNGB3*	SNV (missense)	Original Braunvieh	8.2	001365-9913	[164]
Cholesterol deficiency	*APOB*	retroviral insertion (frameshift)	Holstein	6.1[a]	001965-9913	[80–82, 165]
Hemifacial microsomia	*LAMB1*	SNV (missense)	Romagnola	1.1[b]	002479-9913	[166]
Hypohidrotic ectodermal dysplasia	*EDA*	53kb deletion	Crossbred	*NA*	000543-9913	[167]
Ichthyosis congenita; Syndromic ichthyosis	*FA2H*	1bp insertion (frameshift)	Chianina	5.0[c]	002450-9913	[168]
	DSP	SNV (missense)	Scottish Highland	1.2	002243-9913	[169]
Immunodeficiency with psoriasis-like skin alterations	*IL17RA*	1bp deletion (frameshift)	Holstein	*NA*	002271-9913	[170]
Pulmonary hypoplasia and anasarca syndrome	*ADAMTS3*	SNV (missense)	Cika	*NA*	001562-9913	[171]

AF = allele frequency, OMIA = Online Mendelian Inheritance in Animals [65], ref = references, *NA* = not available

[a]based on 5603 SWISScow genotyped HO animals

[b]based on 221 Romagnola bulls

[c]based on 269 Chianina bulls

As seen from Table 5 and Table 6 the type of identified variants that causes different disorders varies greatly. While SNV can cause lethal phenotypes, it is not surprising that also larger structural variants (insertions and deletions), as well as chromosomal aberrations, cause severe phenotypes.

The study regarding *APOB*-associated cholesterol deficiency in cattle differs from the others, as it does not describe any novel candidate causal variant [165]. Instead, this publication described the co-dominant effects of the variant previously thought to only affect homozygous animals. This has enormous economic impacts for breeders and poses a new animal welfare issue, as with an allele frequency of 6% (based on 5'603 SWISScow genotyped Holstein animals) tens of thousands of cows are potentially affected and can produce affected offspring.

The identification of candidate causative variants becomes more complicated when affected animals present with similar or even identical phenotypes which are caused by different underlying genetic variants, often with differing modes of inheritance, so-called genetic heterogeneity. One example of such an issue is the pulmonary hypoplasia and anasarca syndrome in Cika cattle, this is a population problem caused by a recessively inherited missense variant in the *ADAMTSL3* gene (Table 5); in independent cases of Holstein calves, an identical phenotype arose due to *de novo* trisomy of chromosome 20 (Table 6) [171, 172]. Another complication regarding the detection of candidate variants in the case of recessive disorders is compound heterozygosity. In such cases, our mapping approaches assuming a common genetic background fail. Nevertheless, it was shown that non-parametric linkage can identify disease-associated loci and in combination with WGS, compound heterozygous candidate causative variants can be described [173]. In another example, enough data was available to conduct GWAS, which is hypothesis-free and led to the identification of the associated locus [174].

The opposite issue, phenotypic heterogeneity, describes a situation where a phenotype shows variability in its expression or similar disorders affect the same organ, even though the same genetic variant is responsible. Without knowing the underlying genetic background, it is often unclear if the studied individuals are truly affected by the same genetic disorder complicating the initial study design. An example of this issue could be the two eye disorders achromatopsia and cataract in Original Braunvieh cattle described during my thesis [164, 175]. The main difficulty was to obtain a correct and precise diagnosis from the farmer, which is difficult when only minor symptoms are observed or combinations of disorders occur. If an animal with impaired vision was found, the farmer should have informed us about the origin of it: cataract, blindness or a combination of both. Finally, we were fortunate to find additional calves that had clearly day-blindness, which helped with the description of the observed achromatopsia phenotype and identification of a common recessively inherited variant in the *CNGB3* gene (Table 5) [164]. In addition, we could describe an isolated case of cataract caused by a *de novo* variant in the *ADAMTSL4* gene (Table 6) [175]. Nevertheless, a large part of our

collection of cataract-affected Original Braunvieh animals remains unsolved. Due to the relatedness of these individuals, a recessive inheritance would be assumed; however, as the *ADAMTSL4*-associated cataract case taught us, some of these cases could be due to spontaneous mutation events and affect the locus mapping based on the assumption of a common genetic cause.

These above-described disorders and candidate variants illustrate the genetic and phenotypic complexity of some diseases. Furthermore, it indicates the necessity of precise phenotypic description of the individual cases, preferably by a specialist veterinarian.

Table 6: The herein described dominantly inherited disorders and their associated genes and types of variants identified through the forward genetic approach.

associated disorder	gene	identified variant	breed	OMIA	ref
Achondrogenesis, type II (Bulldog calf syndrome)	*COL2A1*	SNV (missense) 6.7kb deletion 3.4kb deletion	Holstein	001926-9913	[176] [177] [178]
Chondrodysplasia	*FGFR3*	SNV (stop lost)	Holstein	001703-9913	[179]
Classical Ehlers-Danlos syndrome, type II	*COL5A2*	SNV (missense)	Holstein	002295-9913	[180]
Congenital cataract	*ADAMTSL4*	SNV (missense)	Original Braunvieh	002535-9913	[175]
Congenital neuromuscular channelopathy	*KCNG1*	SNV (missense)	Crossbred	002483-9913	[181]
Epidermolysis bullosa simplex	*KRT5*	3bp deletion (in-frame)	Crossbred	002081-9913	[182]
Osteogenesis imperfecta, type II	*COL1A1*	SNV (missense)	Holstein	002127-9913	[183]
Proportional dwarfism with facial dysplasia	many	trisomy 29	Original Braunvieh		[184]
Pulmonary hypoplasia and anasarca syndrome	many	trisomy 20	Holstein	001562-9913	[171, 172]
Skeletal-cardo-enteric dysplasia	*MAP2K2*	SNV (missense)	Romagnola	002381-9913	[185]

OMIA = Online Mendelian Inheritance in Animals [65], ref = references

In the presence of phenotypically well-described spontaneous mutations, such as in our study regarding congenital neuromuscular channelopathy in a Belgian Blue x Holstein crossbred calf, new gene-disease associations and insights into the gene function can be uncovered. We identified for the first time a neurogenic disorder associated with the *KCNG1* gene (Table 6) [181]. In humans, it is often difficult to identify genetic causes of rare diseases due to the limited number of patients. Therefore, this case report can help as a model in human genetics to identify the cause of diseases in patients showing similar symptoms. Livestock animals are predestined as model organisms due to the breeding management system. The use of few sires and AI favors the fast spread of a disorder-causing variant, which can also be quite quickly detected in a larger cohort [89]. In addition, cattle are especially interesting, as the use of genomic information is widespread, pedigree information is recorded attentively, and farmers are encouraged to regularly report any anomalies in newborn animals. Especially, with the use of WGS in single cases potentially affected by *de novo* mutations, novel genetic correlations and findings important for human health can be detected [89]. Lastly, cattle can be useful for the assessment of novel drug treatments, as their organs are larger than in humans, e.g., the eye of a cow has more than double the volume of a human eye.

Assuming a mutation rate of ~1.1×10^{-8} in cattle and a genome size of 2.7Gb, each individual is expected to have approximately 30 *de novo* occurring variants [129]. While most of them are expected to be of a non-coding or synonymous nature, a few are disease-causing. Interestingly, all herein detected dominant *de novo* variants had very severe phenotypes (Table 6).

Concerning the RGA, we managed to identify between 5 and 72 haplotypes using the pgp- and trio-based approach, and 3 to 19 candidate causative variants associated with them per population (Table A8, Figure 11, and Table A9). Most of the haplotypes associated with a high confidence candidate causative variant showed a complete absence of homozygosity (Table 7). These variants are located in genes involved in the cell cycle (e.g., *DIS3* [186]), embryogenesis (e.g., *MRPL55* [187]), or placentation (e.g., *KIR2DS1* [188]), based on their known function in humans and mice. Interestingly, the related populations Brown Swiss and Original Braunvieh showed only a single overlapping haplotype region on chromosome 11 (Figure 11). A clear advantage of the RGA is the possibility to address fertility and rearing success issues without primary knowledge of the underlying biological functions. Thereby, problematics can be addressed, e.g., early embryonic death, that are difficult to observe *in vivo*.

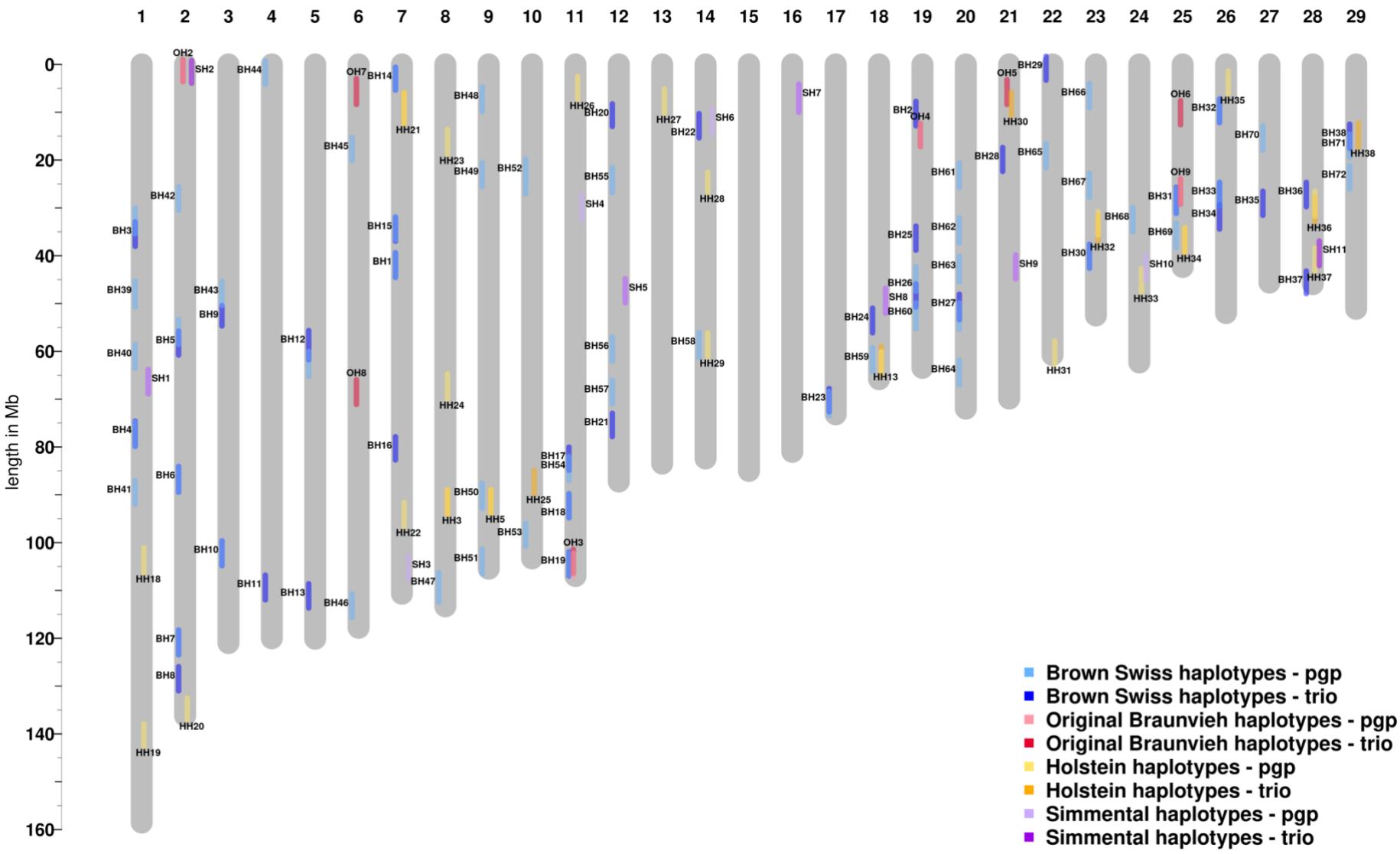


Figure 11: Distribution of all 115 identified haplotype regions showing a deficiency in homozygosity in the four populations Brown Swiss, Original Braunvieh, Holstein, and Simmental across all autosomes.

Table 7: The herein described recessive inherited disorders and their associated genes and types of variants identified through the reverse genetic approach. Note that these are the variants that most likely cause a reduce reproductive success, allowing the identification of the haplotypes showing a depletion in homozygosity.

associated haplotype	gene	identified variant	breed	AF	OMIA	ref
BH6	*MARS2*	SNV (missense)	Brown Swiss	3.4[a]	002517-9913	[189]
BH14	*MRPL55*	SNV (nonsense)	Brown Swiss	3.2[b]	002518-9913	[189]
BH24	*CPT1C*	SNV (missense)	Brown Swiss	2.6[b]	002519-9913	[189]
BH34	*ACSL5*	SNV (missense)	Brown Swiss	2.4[a]	002226-9913	[189]
HH13	*KIR2DS1*	SNV (nonsense)	Holstein	5.0[c]	001836-9913	[190]
HH21	*NOTCH3*	3bp deletion (in-frame deletion)	Holstein	5.5[a]	001908-9913	[190]
HH25	*RIOX1*	30bp deletion (in-frame deletion)	Holstein	2.0[a]	002547-9913	[190]
HH35	*PCDH15*	SNV (missense)	Holstein	2.4[a]	002548-9913	[190]
OH2	*TUBGCP5*	SNV (missense)	Original Braunvieh	4.5[a]	002515-9913	[189]
OH4	*LIG3*	6bp deletion (splice site, frameshift)	Original Braunvieh	3.9[a]	002516-9913	[189]
SH5	*DIS3*	1bp insertion (frameshift)	Simmental	6.2[a]	002505-9913	[191]
SH8	*CYP2B6*	SNV (missense)	Simmental	3.4[a]	002508-9913	[191]
SH9	*NUBPL*	SNV (missense)	Simmental	4.3[a]	002509-9913	[191]

AF = allele frequency, OMIA = Online Mendelian Inheritance in Animals [65], ref = references
[a]no direct genotype available; based on haplotype allele frequency
[b]based on 6577 Brown Swiss SWISScow genotypes
[c]based on 5603 Holstein SWISScow genotypes

I was astonished that many of the haplotypes and GWAS peaks showed a noticeable phenotypic effect. Especially as it can be assumed that many haplotypes are caused by recessive lethal variants. Therefore, one would expect no associations to appear as the effect estimation would be based on heterozygous carriers that should not be affected by a recessive variant. These findings indicate more co-dominant effects of supposedly recessive variants. In order to estimate the dominant effect of the recessive variants, the effects should in future be estimated for non-additive associations [192].

Many haplotypes showing a deficiency in homozygous carrier animals have still unresolved trait associations. Some of them may represent false positives that indicate selection pressure or animal level genotyping bias. Furthermore, for some haplotypes, homozygous carriers were observed in the population. For these cases, an incomplete penetrance can be assumed, which goes against our hypothesis of a recessive lethal variant never occurring in a homozygous state. Therefore, these haplotypes should be further investigated for their underlying phenotypes of both heterozygous and homozygous animals. Instead, complex structural variants, could cause the observed depletion in homozygosity in some genomic regions [193]. Especially deletions and duplications would disrupt our assumption of missing homozygosity. As the expected diploid alleles would actually be haploid or multiploidy, they are falsely predicted in our bioinformatics analysis and cannot be detected. Thus, further statistical analyses (e.g., LD analyses) based on the assumption of diploid organisms would be affected too.

The applied RGA could not be directly transferred to the sex and mitochondrial chromosomes, as they do not show a balanced inheritance from both parents. Therefore, these chromosomes were excluded from the study. Unfortunately, this leads to a lack of knowledge in regard to their contribution to reproduction [194]. The study of Pacheco *et al.* (2020) showed the importance of the X-chromosome on male fertility traits.

Furthermore, association studies performed during this thesis work indicate some genomic regions identified with RGA have a phenotypic impact, but no candidate causative variants could be described yet. Mostly, this is due to the absence of obvious protein-changing variants, and a massive amount of non-coding variation. For non-coding variants, precise prediction of their functional effects was not possible due to the still imperfect annotation status of the bovine genome. Most of the variation we observe is in the non-coding sequence, which includes many evolutionary conserved DNA elements important for gene expression and its regulation; and therefore, should be considered in future studies. Progress on prioritizing functional non-coding variants has been made by integrating evolutionary conservation across species at the position of interest [195]. Such approaches are valuable for prioritizing sequence variants; however, current methods, which would enable us to extract and utilize regulatory sequence information for non-coding variant function prediction, are very limited in bovine genomics. Only recently a study was published listing the branch point sequences that are evolutionary highly conserved in cattle [196]. These are intronic motifs that are essential for proper mRNA splicing [196]. Nevertheless, to carry out a precise allele-specific prediction with single-nucleotide sensitivity, the effects of non-coding variants on transcription factor binding, DNA accessibility, and histone marks of sequences need to

be evaluated systematically. Appropriate resources, which are currently being generated by the Functional Annotation of ANimal Genomes consortium (FAANG), represent a prerequisite to enable systematic evaluation of the predicted regulatory effects of variants in non-coding sequence in its entirety [197].

At present, the understanding of how variable the genome truly is, is widely accepted, especially since the long-read sequencing technologies allow to sequence and analyze several dozen kilo-bases of genome sequence at once [92]. From a technical point of view, the sequencing technology applied is a limiting factor in the detection of SV. In this thesis, 150bp paired-end reads were used and the applied algorithms were able to identify short insertions and deletions up to maximally 50bp reasonably well but longer variants are difficult to identify automatically. This has recently been confirmed in a study comparing different sequencing technologies and workflows, where the prediction accuracy for CNV was the highest for long-read sequencing [198].

Nevertheless, the identification of large SV is possible provided a sufficient number of samples with a well-described diagnosis are available. In a project about cleft palate in piglets, we were able to identify an associated locus by mapping in a large full-sib pedigree using SNP array genotyping data. Subsequently, with WGS data analysis of a case and control animal, we could precisely identify a chromosomal translocation, which was not detected by routine cytogenetic screening [101]. Thereby, we had to screen the locus visually and detected a significant reduction and increase of several Mb on the chromosomes 8 and 14, respectively [101].

A special case of large SV are aneuploidies, where entire chromosomes are duplicated or lost. I was able to identify three cases of aneuploidies in cattle purely based on WGS data (Table 6) [171, 172, 184]. Historically, there are few described cases of aneuploidy in cattle and all of them had been described by applying cytogenetic approaches [199]. With the advent of next-generation sequencing, the cytogenetic approaches might become less important, as with WGS data analyzes we can not only identify trisomies and large SV, but also gain insight into the entire genomic information of an individual at once.

Another limiting factor for the detection of candidate causal variants is the reference sequence used due to gaps and a reference allele bias that leads to worse mapping quality for divergent reads [200]. It was shown that breed-specific reference genome graphs improve read mapping and decrease the reference sequence bias [201]. In cattle, the largest multiassembly graph includes six bovine genomes and includes >70 million

bp more than the ARS-UCD1.2 reference sequence currently used [202]. In-depth analyzes showed that these genomic region that are missing in the reference sequences can have phenotypic associations [203].

Even though this thesis might give the impression that the detection of monogenic disorders is simple, the complexity of certain phenotypes makes it difficult to always pinpoint a single candidate causative variant. Therefore, sometimes there are several potential candidate causative variants that can be associated with the genetic etiology [204, 205]. So far, our prioritization of causative variants was based on their effect prediction as MODIFIER, LOW, MODERATE, and HIGH impact variants. Therein MODERATE and HIGH impact variants are prioritized due to their protein-changing effect. While the effects of LOW impact variants can be interpreted as they are synonymous protein-coding variants, for the intronic and intergenic variants predicted to be MODIFIER we do not know their significance. In a second instance, conservation of the amino acid is evaluated by using across species protein sequence alignment and prediction of deleteriousness of the affected amino acid on the protein function. A novel approach for variant prioritization, besides the variant effect prediction and known gene information, is to evaluate if a variant lies within constrained coding regions [206]. These are genomic regions that were identified by screening massive genomic information to locate regions with reduced protein-changing variation [206]. Similar success rates were found in the search for the genetic causes of rare disorders in humans. A massive study called the 100,000 Genomes Project, recently published preliminary results, where they identified candidate causal variants for 25% of all patients and identified novel gene associations or variants with unknown significance in an additional 10% of cases [207]. This leads to the insight that the focus on monogenic Mendelian disorders is oversimplified and that polygenic constitutions should be considered more frequently.

Further practical issues that keep occurring are the sample acquisition and quality. Usually, farmers simply dispose of the calves that are affected by a severe disorder. Therefore, the availability of cases is very limited, especially for abortions during early embryogenesis, which are rarely recognized. Besides these technical limitations, it should not be forgotten that the topic of this thesis also represents an animal welfare issue. All the described disorders, regardless of the degree of lethality, impair the wellbeing of the affected animals. Especially recessive variants with moderate to high allele frequencies can pose a large risk for the population, e.g., the described *CNGB3*-associated achromatopsia in OB with an allele frequency of 0.082 and the *DIS3*-associated SI haplotype SH5 with an allele frequency of 0.062 (Table 5 and Table 7) [164, 191]. Therefore, it is of the utmost importance to identify the underlying genetic

causes to prevent further spread within populations. This can be achieved by inducing DNA-based selection and avoiding risk matings. I hope that when farmers become more aware of the population-wide impacts and possibilities for improvements, there will be a greater interest in further collaboration with research institutions.

An often-mentioned reason why deleterious alleles keep segregating at a considerable allele frequency is balancing selection [132, 193, 208]. Thereby the disease-associated variant has a desirable effect on production traits and is therefore under positive selection [208]. Within the presented thesis, I did not investigate production traits such as milk yield or composition. Therefore, balancing selection may play a role in the more frequently occurring disease-causing alleles.

In the future, the identified candidate causal variants with recessive inheritance should be considered within the breeding program of respective cattle breeds. With the introduction of the SWISScow genotype array, it is now possible to perform DNA-based precision breeding for these variants. Thereby, over the course of many generations, the presented causal variants could be eliminated from the populations. To manage the risk of a sudden increase in inbreeding, a method such as optimum-contribution selection can be used with a pre-selection against recessive alleles [209].

By using the latest genome editing technologies it would be possible to change the genomic code and eradicate these lethal alleles within a few generations [210]. Johnsson *et al.* (2019) showed with a model that the increased fitness benefit is higher, the more edits that are applied simultaneously, and the more causative variants that are known [210]. Examples, where CRISPR-Cas9-mediated genome editing in cattle have been performed are the introduction of an *MSTN*-disrupting mutation to induce the double muscle phenotype for more efficient meat production [211], and a 3bp deletion in *PMEL* to introduce color dilution in Holstein cattle for better resistance towards heat stress [212]. Earlier technologies, such as the zinc-finger nucleases (ZFN) have been applied in cattle e.g., in the *BLG* gene that is one of the major milk allergens [213], and the transcription activator-like effector nucleases (TALEN) e.g., for the introgression of hornlessness [214]. Thereby, issues can be addressed that concern animal health and welfare, elimination of allergens entering human consumption, disease resistances, and more environmentally friendly production [215].

In conclusion, FGA and RGA have both their advantages and limitations. Nevertheless, both approaches can lead to the detection of causative variants. Regardless, if FGA or RGA was applied, the identified variants described in my thesis impact reproduction, rearing success, or lead to the exclusion of animals from the breeding population due to unwanted phenotypes. Therefore, these most likely pathogenic variants need to be considered within breeding schemes of the concerned populations.

5 Acknowledgments

In 2017, I decided to move to Bern to follow my passionate wish to become a PhD candidate working in the field of genomics in order to improve animal health. Prof. Dr. Cord Drögemüller made this wish come true by inviting me to the Institute of Genetics at the University of Bern. Thank you Cord for your incredible support, understanding, and great mentoring.

This thesis was only possible thanks to the support of many farmers and industrial collaborators. These included partners such as the breeding associations, artificial insemination companies, and the Genomic Selection team of Qualitas AG. I want to mention Franz Seefried especially, he was one of the driving forces in this PhD project. Thank you, Franz, for your great support and patience.

Many of the presented studies were supported by national and international researchers. I am very grateful to all of them for their openness and trust to collaborate. All your inputs were very inspiring and made working on the individual projects compelling.

Furthermore, I want to thank the lab technicians, especially Nathalie Besuchet, who were incredibly helpful and vital for the success of this thesis. They prepared hundreds of DNA samples for diverse analyzes and conducted genotyping whenever necessary.

Likewise, I want to acknowledge the thesis committee. Dr. Rémy Bruggmann for being my co-supervisor, Prof. Pierre Meister for being my mentor, and Dr. Aurélien Capitan for agreeing to be my external co-referee. Thank you all for your valuable inputs regarding different aspects of my thesis.

It is dear to my heart to thank the whole team at the Institute of Genetics. You were a great and inspiring support during my time here. I very much enjoyed working with you all and problem solving, including all the brilliant/crazy out-of-the-box ideas. I especially want to thank my fellow musketeers Anna Letko and Jan Henkel. You guys made this journey unforgettable.

I want to thank the people who proof-read the presented thesis. Thank you very much for your contributions.

Generally, I want to thank all my mentors. It does not matter if we met only briefly, only for a certain time, or are in touch daily. You guide me and lend me a helping hand whenever I am in need of support and therefore, I am very thankful. I want to mention by name, Britta Lundström who through the VetMent program mentored me officially during this turbulent time, called PhD.

I want to thank my family and friends. Thank you all very much for your support and faith in me. To my local people: I am glad you managed to get me out of the office from time to time. To my international people: You keep inspiring me to look outside the box and think about the bigger picture. I am looking forward to visiting you on my further journeys.

“Legst de best!” Finally, I thank my partner, Roger Dubach. Thank you for always believing in me and supporting me even during the craziest of times. I am looking forward to more fabulous adventures with you.

6 Curriculum vitae

Removed due to data privacy reasons.

7 List of publications

7.1 Publications in peer-reviewed scientific journals

7.1.1 First author publications

Häfliger I.M., Agerholm J.S., and Drögemüller C. **2020**. Constitutional trisomy 20 in an aborted Holstein fetus with pulmonary hypoplasia and anasarca syndrome. *Animal Genetics.* DOI: 10.1111/age.13008

Häfliger I.M., Behn H., Freick M., Jagannathan V., and Drögemüller C. **2019**. A *COL2A1 de novo* variant in a Holstein bulldog calf. *Animal Genetics*. DOI: 10.1111/age.12735

Häfliger I.M., Hirter N., Paris J., and Drögemüller C. **2020**. A *de novo* germline mutation of *KIT* in a white-spotted Brown Swiss cow. *Animal Genetics*. DOI: 10.1111/age.12920

Häfliger I.M., Hofstetter S., Mock T., Stettler M.H., Meylan M., Mehinagic K., Stokar-Regenscheit N., and Drögemüller C. **2019**. *APOB*-associated cholesterol deficiency in Holstein cattle is not a simple recessive disease. *Animal Genetics*. DOI: 10.1111/age.12801

Häfliger I.M., Koch C.T., Michel A., Rüfenacht S., Meylan M., Welle M.M., and Drögemüller C. **2021**. *DSP* missense variant in a Scottish Highland calf with congenital ichthyosis, alopecia, acantholysis of the tongue and corneal defects. *BMC Veterinary research.* DOI: 10.1186/s12917-021-03113-3

Häfliger I.M., Letko A., Murgiano L., and Drögemüller C. **2020**. *De novo* stop-lost germline mutation in *FGFR3* causes severe chondrodysplasia in the progeny of a Holstein bull. *Animal Genetics.* DOI: 10.1111/age.12934

Häfliger I.M.[*], Marchionatti E.*, Stengård M.*, Wolf-Hofstetter S., Paris J.M., Jacinto G.P.J., Watté C., Voelter K., Occelli L.M., Komáromy A., Oevermann A., Goepfert C., Borgo A., Roduit R., Spengeler M., Seefreid F., and Drögemüller C. **2021**. *CNGB3* missense variant causes recessive achromatopsia in Original Braunvieh cattle. *International Journal of Molecular Sciences*. DOI: 10.3390/ijms222212440.

Häfliger I.M., Seefried F., and Drögemüller C. **2020.** Trisomy 29 in a stillborn Swiss Original Braunvieh calf. *Animal Genetics*. DOI: 10.1111/age.12929

Häfliger I.M., Seefried F., and Drögemüller C. **2021**. Reverse genetic screen for deleterious recessive variants in the local Simmental cattle population of Switzerland. *Animals.* DOI: 10.3390/ani11123535

Häfliger I.M., Seefried F., Spengeler M., and Drögemüller C. **2021**. Mining massive genomic data of two Swiss Braunvieh cattle populations reveals six novel candidate variants that impair reproductive success. *Genetics Selection Evolution.* DOI: 10.1186/s12711-021-00686-3

Häfliger I.M.*, Sickinger M.*, Holsteg M., Raeder L.M., Henrich M., Drögemüller C., and Lühken G. **2020**. An *IL17RA* frameshift variant in a Holstein cattle family with psoriasis-

[*] equally contributing authors

like skin alterations and immunodeficiency. *BMC Genetics.* DOI: 10.1186/s12863-020-00860-4

Häfliger I.M., Spengeler M., Seefried F., and Drögemüller C. **2022**. Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle. *Scientific Reports*. DOI: 10.1038/s41598-022-09403-6

Häfliger I.M., Wiedemar N., Švara T., Starič J., Cociancich V., Šest K., Gombač M., Paller T., Agerholm J.S., and Drögemüller C. **2020**. Identification of small and large genomic variants in bovine pulmonary hypoplasia and anasarca syndrome. *Animal Genetics*. DOI: 10.1111/age.12923

Häfliger I.M., Wolf-Hofstetter S., Casola C., Hetzel U., Seefried F.R., and Drögemüller C. **2022**. A *de novo* mutation in the bovine *ADAMTSL4* gene in an Original Braunvieh calf with congenital cataract. *Animal Genetics*. DOI: 10.1111/age.13178

Jacinto J.G.P.*, **Häfliger I.M.***, Veiga I.M.B.*, Letko A., Benazzi C., Bolcato M., and Drögemüller C. **2020**. A heterozygous missense variant in the *COL5A2* in Holstein cattle resembling the classical Ehlers-Danlos syndrome. *Animals*. DOI: 10.3390/ani10112002

7.1.2 Co-author publications

Brunner M.A.T., Veiga I., Niggeler A., **Häfliger I.M.**, Stettler M., Meylan M., Welle M., and Drögemüller C. **2020**. Is a *de novo* nonsense variant in the *ASPDH* gene the cause for ulcerative skin lesions in a Holstein calf?. *Veterinary Dermatology*. DOI: 10.1111/vde.12827

Butty A.M., Chud T.C.S., Cardoso D.F., Lopes L.S.F., Miglior F., Schenkel F.S., Cánovas A., **Häfliger I.M.**, Drögemüller C., Stothard P., Malchiodi F., and Baes C.F. **2021**. Genome-wide association study between copy number variants and hoof health traits in Holstein dairy cattle. *Journal of Dairy Science*. DOI: 10.3168/jds.2020-19879

Butty A.M., Chud T.C.S., Miglior F., Schenkel F.S., Kommadath A., Krivushin K., Grant J.R., **Häfliger I.M.**, Drögemüller C., Cánovas A., Stothard P., and Baes C.F. **2019.** High confidence copy number variants identified in Holstein dairy cattle from whole genome sequence and genotype array data. *Scientific Reports*. DOI: 10.1038/s41598-020-64680-3

Eager K.L.M., Cauchi M., Willet C.E., **Häfliger I.M.**, Drögemüller C., O'Rourke B.A., and. Tammen I. **2021.** The previously reported *LRP4* c.4940C>T variant is not associated with syndactyly in cattle. *Animal Genetics,* 2021 Mar 22. DOI: 10.1111/age.13061

Grahofer A., Letko A., **Häfliger I.M.**, Jagannathan V., Ducos A., Richard O., Peter V., Nathues H., and Drögemüller C. **2019**. Chromosomal imbalance in pigs showing a syndromic form of cleft palate. *BMC Genomics*. DOI: 10.1186/s12864-019-5711-4

* equally contributing authors

Hirter N., Letko A., **Häfliger I.M.**, Becker D., Greber D., and Drögemüller C. **2020.** A genome-wide significant association on chromosome 15 for congenital entropion in Swiss White Alpine sheep. *Animal Genetics.* DOI: 10.1111/age.12903

Hofstetter S., Seefried F., **Häfliger I.M.**, Jagannathan V., Leeb T., and Drögemüller C. **2018**. A non-coding regulatory variant in the 5'-region of the *MITF* gene is associated with white spotted coat in Brown Swiss cattle. *Animal Genetics*. DOI: 10.1111/age.12751

Jacinto J.G.P., **Häfliger I.M.**, Akyürek E.E., Sacchetto R., Benazzi C., Gentile A., and Drögemüller C. **2021**. *KCNG1*-related syndromic form of congenital neuromuscular channelopathy in a crossbred calf. *Genes*. DOI: 10.3390/genes12111792

Jacinto J.G.P., **Häfliger I.M.**, Bernardini M., Mandara M., Bianchi E., Bolcato M., Romagnoli N., Gentile A., and Drögemüller C. **2021.** A homozygous missense variant in laminin subunit beta 1 as candidate causal mutation of hemifacial microsomia in Romagnola cattle. *Journal of Veterinary Internal Medicine*. DOI: 10.1111/jvim.16316

Jacinto J.G.P., **Häfliger I.M.**, Borel N., Zanolari P., Drögemüller C. and Veiga I.M.B. **2021.** Clinicopathological and Genomic Characterization of a Simmental Calf with Generalized Bovine Juvenile Angiomatosis. *Animals*. DOI: 10.3390/ani11030624

Jacinto J.G.P., **Häfliger I.M.**, Gentile A., and Drögemüller C. **2021.** A heterozygous missense variant in *MAP2K2* in a stillborn Romagnola calf with skeletal-cardio-enteric dysplasia. *Animals*. DOI: 10.3390/ani11071931

Jacinto J.G.P., **Häfliger I.M.**, Gentile A., and Drögemüller C. **2020**. A 6.7kb deletion in the *COL2A1* gene in a Holstein calf with achondrogenesis type II and perosomus elumbis. *Animal Genetics*. DOI: 10.1111/age.13033

Jacinto J.G.P., **Häfliger I.M.**, McEvoy F., Drögemüller C., and Agerholm J.S. **2021.** A *de novo* mutation in *COL1A1* in a Holstein calf with osteo-genesis imperfecta type II. *Animals*. DOI: 10.3390/ani11020561

Jacinto J.G.P., **Häfliger I.M.**, Veiga M.B., Drögemüller C., and Agerholm J.S. **2020**. A *de novo* mutation in *KRT5* in a crossbred calf with epidermolysis bullosa. *Journal of Veterinary Internal Medicine*. DOI: 10.1111/jvim.15943

Jacinto J.G.P., **Häfliger I.M.**, Veiga I.M.B., Letko A., Gentile A., and Drögemüller C. **2021**. A frameshift insertion in *FA2H* causes a recessively inherited form of ichthyosis congenita in Chianina cattle. *Molecular Genetics and Genomics*. DOI: 10.1007/s00438-021-01824-8

Joller S., **Häfliger I.M.**, Drögemüller C., Richard O.K., and Grahofer A. **2020**. Thrombocytopenic purpura on an organic farm with pen mating: a case report on the re-emergence of an old disease. *Porcine Health Management*. DOI: 10.1186/s40813-020-00157-z

Küttel L., Letko A., **Häfliger, I.M.**, Signer-Hasler H., Joller S., Hirsbrunner G., Mészáros G., Sölkner J., Flury C., Leeb T., and Drögemüller C. **2019**. A complex structural variant at the *KIT* locus in cattle with the Pinzgauer spotting pattern. *Animal Genetics*. DOI: 10.1111/age.12821

Letko A., Dijkman R., Strugnell B., **Häfliger I.M.**, Paris J.M., Henderson K., Geraghty T., Orr H., Scholes S., and Drögemüller C. **2020**. Deleterious *AGXT* missense variant

associated with type 1 primary hyperoxaluria (PH1) in Zwartbles sheep. *Genes*. DOI: 10.3390/genes11101147

Letko A., Strugnell B., **Häfliger I.M.**, Paris J.M., Waine K., Drögemüller C., and Scholes S. **2020**. Compound heterozygous *PLA2G6* loss-of-function variants in Swaledale sheep with neuroaxonal dystrophy. *Molecular Genetics and Genomics*. DOI: 10.1007/s00438-020-01742-1

Mock T., Michel B., Dettwiler M., Rodriguez-Campos S., **Häfliger I.M.**, Drögemüller C., Mee J.F., Hüsler J., Bodmer M. and Hirsbrunner G. **2020**. Evaluation of an investigative model in dairy herds with high calf perinatal mortality rates in Switzerland. *Theriogenology*. DOI: 10.1016/j.theriogenology.2020.02.039

O'Toole D., **Häfliger I.M.**, Leuthard F., Schumaker B., Steadman L., Murphy B., Drögemüller C., and Leeb T. **2021**. X-linked hypohidrotic ectodermal dysplasia in crossbred beef cattle due to a large deletion in *EDA*. *Animals*. DOI: 10.3390/ani11030657

Paris J.M., Letko A., **Häfliger I.M.**, Ammann P., and Drögemüller C. 2020. Ear type in sheep is associated with the *MSRB3* locus. *Animal Genetics*. DOI: 10.1111/age.12994

Paris J.M., Letko A., **Häfliger I.M.**, Ammann P., Flury C., and Drögemüller C. **2019**. Identification of two *TYRP1* loss-of-function alleles in Valais Red sheep. *Animal Genetics*. DOI: 10.1111/age.12863

Paris J., Letko A., **Häfliger I.M.**, Gombač M., Klinc P., Škibin A., Pogorevc E., Švara T., and Drögemüller C. **2020**. A *de novo* variant in *OTX2* in a sheep with otocephaly. *Acta Veterinaria Scandinavica*. DOI: 10.1186/s13028-020-0503-z

Simon R., Lischer H., Pieńkowska-Schelling A., Keller I., **Häfliger I.M.**, Letko A., Schelling C., Lühken G., and Drögemüller C. **2020**. New genomic features of the polled intersex syndrome variant in goats unraveled by long-read whole-genome sequencing. *Animal Genetics*. DOI: 10.1111/age.12918

Widmer S., Seefried F.R., von Rohr P., **Häfliger I.M.**, Spengeler M., and Drögemüller C. **2021**. A major QTL at *LHCGR* for multiple birth in Holstein cattle. *Genetics Selection Evolution*. DOI: 10.1186/s12711-021-00650-1

Woolley S.A., Eager K., **Häfliger I.M.**, Bauer A., Drögemüller C., Leeb T., O'Rourke B.A., and Tammen I. **2019**. An *ABCA12* missense variant in a Shorthorn calf with ichthyosis fetalis. *Animal Genetics*. DOI: 10.1111/age.12856

7.2 Conference attendances and invited talks

7.2.1 Oral presentations at conferences

Häfliger I.M., Seefried F.R., and Drögemüller C. (2021). Entwicklung eines SNP-chips für die Schweizer Rinderzucht. Netzwerk Nutztiere Vortragstagung, Grangeneuve, CH, Nov 11 **2021**

Häfliger I.M., Seefried F.R., and Drögemüller C. (2021). Numerous monogenic recessive disorders segregate in Swiss dairy populations. Swiss Animal Breeding Technology Platform (SABRE-TP), Zurich, November 10 **2021**.

Häfliger I.M., Marchionatti E., Wolf-Hofstetter S., Paris J.M., Jacinto J.G.P., Watté C., Stengard M., Voelter K., Occelli L.M., Komáromy A.M., Oevermann A., Göpfert C., Borgo A., Roduit R., Seefried F., and Drögemüller C. (2021). Missense variant in *CNGB3* causes recessive day-blindness (achromatopsia) in Original Braunvieh cattle. 72nd Annual Meeting of the European Federation of Animal Science, hybrid meeting: virtual and Davos, CH, Aug 30 - Sept 3 **2021**

Häfliger I.M., Seefried F.R., and Drögemüller C. (2021). Exploiting massive genotyping and sequencing data to improve fertility in Swiss dairy. 72nd Annual Meeting of the European Federation of Animal Science, hybrid meeting: virtual and Davos, CH, Aug 30 - Sept 3 **2021**

Häfliger I.M., Seefried F.R., and Drögemüller C. (2019). Exploiting massive genomic data to improve fertility and rearing success in Swiss dairy cattle. 17th International Conference on Production Disease in Farm Animals, Bern, CH, June 28 **2019**

7.2.2 Poster presentations at conferences

Häfliger I.M., Sickinger S., Holsteg M., Raeder L.M., Henrich M., Marquardt S., Drögemüller C., and Lühken G. (2020). A recessive *IL17RA* frameshift variant in Holstein cattle with psoriasis-like skin alterations and immunodeficiency. American Dairy Science Association Annual Meeting, virtual, June 22-24 **2020**

Häfliger I.M., Agerholm J.S., and Drögemüller C. (2019). A case of pulmonary hypoplasia and anasarca syndrome in Holstein cattle due to trisomy of chromosome 20. 37th Conference of the International Society for Animal Genetics, Lleida, ES, July 7-12 **2019**

Häfliger I.M., Seefried F.R., and Drögemüller C. (2019). A comprehensive genetic screen for recessive variants impairing fertility and rearing success in the Swiss Original Braunvieh Cattle. Symposium of the Swiss Association of Animal Sciences, Lindau, CH, April 16 **2019**

7.2.3 Invited talks

Ergebnisse aus dem Missing Homozygosity Projekt – HOL. Forschungsausschuss Association Swiss cattle breeders (ASR), online, November 5 **2021.**

Missing homozygosity in Swiss Braunvieh cattle. Department of Veterinary Medical Sciences, University of Bologna, online, June 9 **2021.**

Ergebnisse aus dem Missing Homozygosity Projekt – BSW. Forschungsausschuss Association Swiss cattle breeders (ASR), online, May 21 **2021.**

A comprehensive genetic screen for recessive variants impairing fertility and rearing success of Swiss cattle. Lab meeting, Ruminants Clinics, Vetsuisse Faculty, University of Bern, Online, April 23 **2021.**

Development of gene tests and management of hereditary defects. Seminar of the Grey cattle breeder, Nauders, AT, Jan. 11 **2019.**

A comprehensive genetic screen for recessive mutations impairing fertility and rearing success of Swiss cattle. Institute internal meeting, Center for Genetic Improvement of Livestock, University of Guelph, CA, December 12 **2018**

8 Appendix

Table A1: Overview of the traits in Swiss breeding programs that concern traits related to fertility, birth, growth and survival.

trait group	trait	description
fertility	non-return rate heifer	heifers non-return rate after 56 days, binary
	non-return rate cow	cows non-return rate after 56 days, binary
	interval first to last insemination heifer	interval between first and last insemination for heifer, days
	interval first to last insemination cow	interval between first and last insemination for cows, days
	interval calving to insemination	interval from calving to first service, days
birth history direct	percentage normal births	calving ease, scored between 1-without help to 5-dystocia
	percentage live births	percentage of calves born alive
	birth weight	weight of calve at birth, kg
	gestation length	days from successful insemination to birth
	multiple birth	percentage of multiple births
birth history maternal	percentage normal births	calving ease, scored between 1-without help to 5-dystocia
	percentage live births	percentage of calves born alive
	birth weight	weight of calve at birth, kg
	gestation length	days from successful insemination to birth
	multiple birth	percentage of multiple births
rearing success	survival period 1	survival from day 3 up to 30th day of life
	survival heifer period 2	survival of heifers from day 31 up to 458 days
	survival bull period 2	survival of young bulls from 31 days up to 183 days
slaughter traits calves	slaughter weight	weight at slaughter, kg
	carcass conformation score	amount of meat at slaughter, kg
	carcass fat score	fat cover in the meat
slaughter traits adults	slaughter weight	weight at slaughter, kg
	carcass conformation score	amount of meat at slaughter, kg
	carcass fat score	fat cover in the meat

Table A2: Skin disorders described in cattle with their associated genes and types of variants.

	MOI	gene	OMIA	type of variant	breed	ref
Ichthyosis	AR	*ABCA12*	002238-9913	SNV (missense, frameshift)	Chianina, Shorthorn, Polled Hereford	[216–218]
Dystrophic epidermolysis bullosa	AR	*COL7A1*	000341-9913	SNV (nonsense)	Rotes Höhenvieh, Vorderwald	[219, 220]
Tricho-dento-osseous-like syndrome	AD*	*DLX3*	002109-9913	10bp insertion (frameshift)	Brown Swiss	[221]
Ectodermal dysplasia	XR	*EDA*	000543-9913	gross deletion; gross inversion; 161bp insertion, 19bp deletion, 4bp insertion (frameshift); SNV (splice site, nonsense)	Holstein, Japanese Black, crossbred	[222–230]
Anhidrotic ectodermal dysplasia	AR	*EDAR*	002128-9913	1bp insertion (frameshift)	Charolais	[89]
Ehlers-Danlos syndrome	AR	*EPYC*	001716-9913	SNV (missense)	Holstein	[231]
Hypotrichosis	AR	*HEPHL1*	002230-9913	SNV (nonsense)	Belted Galloway	[232]
Epidermolysis bullosa, junctionalis	AR	*ITGB4*	001948-9913	gross deletion	Charolais	[233, 234]
Epidermolysis bullosa simplex	AD*	*KRT5*	002081-9913	SNV (missense)	Friesian X Jersey crossbred	[235]
Hypotrichosis	AR	*KRT71*	002114-9913	8bp deletion (frameshift)	Hereford	[236]
Epidermolysis bullosa junctionalis	AR	*LAMA3*	001677-9913	SNV (nonsense)	Belgian Blue	[237]
Epidermolysis bullosa junctionalis	AR	*LAMC2*	001678-9913	gross deletion	Hereford	[238]
Hypotrichosis with coat-color dilution	AD	*MC1R*	001544-9913	SNV (missense)	crossbred	[239]
Zinc deficiency-like syndrome (genodermatosis)	AR	*PLD4*	001935-9913	SNV (nonsense)	Fleckvieh	[240]

Hypotrichosis with coat-color dilution	AD	*PMEL*	001544-9913	3bp deletion (in-frame)	crossbred	[239]
Coat color, albinism, oculocutaneous type IV	AR	*SLC45A2*	001821-9913	SNV (missense)	Brown Swiss	[78]
Streaked hypotrichosis	XD	*TSR2*	000542-9913	SNV (splice site)	Pezzata Rosso	[241]
Coat color, oculocutaneous albinism type I	AR	*TYR*	000202-9913	1bp insertion (frameshift)	Brown Swiss	[77]

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, XR = X-linked recessive, XD = X-linked dominant, OMIA = Online Mendelian Inheritance in Animals [65], ref = references

**de novo* mutation

Table A3: Bone morphology altering disorders described in cattle with their associated genes and types of variants.

phenotype	MOI	gene	OMIA	type of variant	breed	ref
Dwarfism (Bulldog calf)	AR	*ACAN*	001271-9913	4bp insertion (frameshift); 1bp insertion (frameshift); SNV (regulatory)	Dexter, Scottish Highland, Zebu	[71, 242]
Tibial hemimelia	AR	*ALX4*	001009-9913	gross deletion; 20bp duplication (frameshift)	Shorthorn, Galloway	[243, 244]
Caprine-like generalized hypoplasia syndrome	AR	*CEP250*	001502-9913	SNV (nonsense)	Montbeliarde	[245]
Osteogenesis imperfecta type II	AD*	*COL1A1*	002127-9913	4bp deletion + 1bp insertion (in-frame deletion + missense); SNV (missense)	Fleckvieh, Red Angus	[89, 246]
Achondrogenesis type II	AD*	*COL2A1*	001926-9913	SNV (splice site, missense)	Holstein, crossbred	[89, 144, 247, 248]
Mandibulofacial dysostosis	AR	*CYP26C1*	002288-9913	SNV (missense)	Hereford	[249]
Brachyspina	AR	*FANCI*	000151-9913	gross deletion	Holstein	[250, 251]
Marfan syndrome	AD*	*FBN1*	000628-9913	SNV (missense, splice site)	Limousin, Japanese Black	[252, 253]
Facial dysplasia syndrome	AD*	*FGFR2*	002090-9913	SNV (missense)	Holstein	[254]
Dwarfism	AR	*GON4L*	001985-9913	1bp deletion (frameshift)	Fleckvieh	[255]
Arachnomelia	AR	*MOCS1*	001541-9913	2bp deletion (frameshift)	Fleckvieh	[256, 257]
Crooked tail	AR	*MRC2*	001452-9913	2bp deletion (frameshift); SNV (missense)	Belgian Blue	[258, 259]
Dwarfism	AR	*PRKG2*	001485-9913	SNV (nonsense)	Angus	[260]
Proportionate dwarfism with inflammatory lesions	AR	*RNF11*	001686-9913	SNV (splice site)	Belgian Blue	[261]
Osteopetrosis	AR	*SLC4A2*	002443-9913	3.8kb deletion	Angus, Holstein, Hereford, Simmental	[262]
Complex vertebral malformation (CVM)	AR	*SLC35A3*	001340-9913	SNV (missense)	Holstein	[263, 264]
Acrodermatitis enteropathica	AR	*SLC39A4*	000593-9913	SNV (splice site)	Holstein	[265]

Arachnomelia	AR	*SUOX*	000059-9913	1bp insertion (frameshift)	Brown Swiss	[266]
Vertebral and spinal dysplasia	AD[&]	*T*	001951-9913	SNV (missense)	Holstein	[267]
Frontonasal dysplasia	AD*	*ZIC2*	002307-9913	1bp deletion (frameshift)	Limousin	[268]

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, OMIA = Online Mendelian Inheritance in Animals [65], ref = references

**de novo* mutation

[&]incomplete penetrance

Table A4: Neuromuscular disorders described in cattle with their associated genes and types of variants.

phenotype	MOI	gene	OMIA	type of variant	breed	ref
Contractural arachnodactyly (Fawn calf syndrome)	AR	*ADAMTSL3*	001511-9913	gross deletion	Angus	[269]
Arthrogryposis multiplex congenita	AR	*AGRN*	002135-9913	gross deletion	Angus	[270]
Citrullinaemia	AR	*ASS1*	000194-9913	SNV (nonsense)	Holstein	[271]
Congenital muscular dystonia 1	AR	*ATP2A1*	001450-9913	SNV (missense)	Belgian Blue	[218]
Congenital pseudomyotonia	AR	*ATP2A1*	001464-9913	SNV (missense)	Chianina, Romagnola	[272, 273]
Maple syrup urine disease	AR	*BCKDHA*	000627-9913	SNV (missense, nonsense)	Hereford, Shorthorn	[274, 275]
Neurocristopathy	AD*	*CHD7*	002125-9913	5bp deletion (frameshift)	Montbeliard	[89]
Arthrogryposis multiplex	AR	*CHRNB1*	002022-9913	1bp deletion (frameshift)	Red Danish	[276]
Congenital myasthenic syndrome	AR	*CHRNE*	000685-9913	20bp deletion (frameshift)	Brahman	[277]
Neuronal ceroid lipofuscinosis 5	AR	*CLN5*	001482-9913	1bp duplication (frameshift)	Devon	[278]
Charcot Marie Tooth disease	AR	*FGD4*	002374-9913	SNV (splice site)	Holstein, Jersey	[192]
Glycogen storage disease II	AR	*GAA*	000419-9913	SNV (missense, nonsense)	Shorthorn, Brahman, Droughtmaster	[279]
Myoclonus	AR	*GLRA1*	000689-9913	SNV (nonsense)	Hereford	[280]
Spinal muscular atrophy (SMA)	AR	*KDSR (FVT1)*	002390-9913	SNV (missense)	Brown Swiss	[281, 282]
Congenityl hypomyelinogenesis	AR	*KIF1C*	000527-9913	SNV (missense)	Charolais	[283]
Congenital muscular dystonia 2	AR	*LOC528050*	001451-9913	SNV (missense)	Belgian Blue	[218]
Mannosidosis, beta	AR	*MANBA*	000626-9913	SNV (nonsense)	Salers	[284]
Mannosidosis, alpha	AR	*MAN2B1*	000625-9913	SNV (missense)	Galloway, Angus, Murray Grey	[285]
Degenerative axonopathy	AR	*MFN2*	001106-9913	SNV (splice site)	Tyrolean Grey	[286]
Distal arthrogryposis type 1B	AD*	*MYBPC1*	001978-9913	SNV (missense)	Holstein	[287]
Mucopolysaccharidosis IIIB	AR	*NAGLU*	001342-9913	SNV (missense)	unknown	[288]

Niemann-Pick disease, type C1	AR	*NPC1*	000725-9913	SNV (missense)	Angus	[289]
Arthrogryposis, lethal syndrome	AR	*PIGH*	001953-9913	SNV (splice site)	Belgian Blue	[290]
Myeloencephalopathy (Weaver syndrome)	AR	*PNPLA8*	000827-9913	SNV (missense)	Brown Swiss	[291–293]
Susceptibility/resistance to spongiform encephalopathy	MF	*PRNP*	000944-9913	SNV (missense)	numerous	[294]
Glycogen storage disease V	AR	*PYGM*	001139-9913	SNV (missense)	Charolais	[295]
Tetradysmelia	AR	*RSPO2*	002297-9913	gross deletion (frameshift)	Holstein	[296]
Turning calves syndrome	AR	*SLC25A46*	002150-9913	SNV (missense)	Rouge-des-Prés	[297]
Spinal dysmyelination	AR	*SPAST*	001247-9913	SNV (missense)	Brown Swiss	[298, 299]
Neuropathy with splayed forelimbs	AR	*UCHL1*	002298-9913	SNV (missense)	Jersey	[300]

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, MF = multifactorial, OMIA = Online Mendelian Inheritance in Animals [65], ref = references

**de novo* mutation

Table A5: Eye disorders described in cattle with their associated genes and types of variants.

phenotype	MOI	gene	OMIA	type of variant	breed	ref
Cataract	AR	*CPAMD8*	002111-9913	SNV (nonsense)	Holstein	[75]
Cataract	AR	*NID1*	001936-9913	855bp deletion (frameshift)	Romagnola	[74]
Retinitis pigmentosa 1	AR	*RP1*	002029-9913	1bp insertion (frameshift)	Angus, Beef Booster Composite, Belgian Blue, Charolais, Gelbvieh, Holstein, Maine Anjou, Normande, Red Angus	[76]
Coat color, albinism, oculocutaneous type IV	AR	*SLC45A2*	001821-9913	SNV (missense)	Brown Swiss	[78]
Coat color, oculocutaneous albinism type I	AR	*TYR*	000202-9913	1bp insertion (frameshift)	Brown Swiss	[77]

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, OMIA = Online Mendelian Inheritance in Animals [65], ref = references

Table A6: Known genes and type of variants affecting fertility and rearing success, including fatal syndromic disorders in cattle.

phenotype	MOI	gene	OMIA	type of variant	breed	ref
Male fertility	AR	*ABHD16B*		SNV (missense)	Holstein	[301]
Cholesterol deficiency (CDH)	AR	*APOB*	001965-9913	1.3kb ERV insertion (frameshift)	Holstein	[79, 80, 82]
Asthenospermia	AR	*CCDC189*	002167-9913	SNV (splice site)	Nordic Red	[302]
Renal dysplasia	AR	*CLDN16*	001135-9913	gross deletion	Japanese Black	[303, 304]
Protoporphyria	MF	*FECH*	000836-9913	SNV (stop lost)	Blond d'Aquitaine, Limousin	[305]
Haemophilia A	XR	*F8*	000437-9913	SNV (missense)	Japanese Brown, Fleckvieh	[306, 307]
Factor XI deficiency (blood coagulation)	AR	*F11*	000363-9913	76bp insertion (frameshift); 15bp insertion (missense + in-frame)	Holstein, Sahiwal, Japanese Black	[308, 309]
Congenital disorder of glycosylation	AR	*GALNT2*	002375-9913	SNV (splice site)	Holstein, Jersey	[192]
Crop ears	AD[&]	*HMX1*	000317-9913	76bp duplication	Highland cattle	[310]
Hereditary perinatal weak calf syndrome	AR	*IARS*	001817-9913	SNV (missense)d	Japanese Black	[311]
Bovine leukocyte adhesion deficiency (BLAD)	AR	*ITGB2*	000595-9913	SNV (missense)	Holstein	[312]
Lethal multi-organ developmental dysplasia (Paunch calf syndrome)	AR	*KDM2B*	001722-9913	SNV (missense)	Marchigiana; Romagnola	[313]
Gonadal hypoplasia	AR	*KIT*	000426-9913	structural rearrangement	Northern Finncattle; Swedish Mountain	[314]
Syndactyly (mule foot)	AR	*LRP4*	000963-9913	SNV (missense, splice site); 2bp deletion + insertion (2bp missense)	Holstein; Simmental; Angus	[315–318]
Chediak-Higashi syndrome	AR	*LYST*	000185-9913	SNV (missense)	Japanese Black	[319]
Abortion and stillbirth	AD*	*MIMT1*	001565-9913	gross deletion	Finnish Ayrshire	[320]
Coat colour, white spotting	MF	*MITF*	000214-9913	SNV (regulatory)	Holstein, Simmental, Brown Swiss	[321, 322]

Depigmentation associated with microphthalmia	AD*	*MITF*	001931-9913	gross deletion	Holstein	[323]
Dominant white with bilateral deafness	AD*	*MITF*	001680-9913	3bp deletion (in-frame); SNV (missense)	Holstein, Fleckvieh	[89, 324]
Xanthinuria type II	AR	*MOCOS*	001819-9913	1bp deletion (frameshift)	Tyrolean Grey; Original Braunvieh	[35, 325]
Muscular hypertrophy (double muscling)	AD[&]	*MSTN*	000683-9913	11bp deletion, 7bp deletion + 11bp insertion (frameshift); SNV (missense, nonsense)	Angus, Asturiana, Belgian Blue, Blonde d'Aquitaine, Braford, Charolais, Gasconne, Gelbvieh, Limousin, Maine Anjou, Marchigiana, Murray Grey, Parthenaise, Piedmontese, Santa Gertrudis, South Devon	[326–331]
Developmental duplications	AR	*NHLRC2*	002103-9913	SNV (missense)	Angus	[332]
Dilated cardyomyopathy	AR	*OPA3*	000162-9913	SNV (missense)	Fleckvieh	[333]
Lethal arthrogryposis syndrome	AR	*PIGH*	001953-9913	SNV (splice site)	Belgian Blue	[290]
Cardiomyopathy and wolly haircoat syndrome	AR	*PPP1R13L*	000161-9913	7bp duplication (frameshift)	Hereford	[334]
Hairy	AD	*PRL*	000441-9913	SNV (missense)	Fleckvieh, Holstein, Jersey	[335]
Morphological abnormalities in sperm and male fertility	AR	*QRICH2*		1bp deletion (frameshift)	Brown Swiss	[52]
Thrombopathia	AR	*RASGRP2*	002433-9913	SNV (missense)	Fleckvieh, Simmental	[336]
Slick hair	AD	*PRLR*	001372-9913	SNV (nonsense, frameshift)	Carora, Limonero, Romosinuano, Senepol	[335, 337]
Embryonic death	AR	*RNASEH2B*	001901-9913	gross deletion	Nordic Holstein	[193]
Male fertility	AR	*SPATA16*		SNV (missense)	Brown Swiss	[51]
Male subfertility	AR	*TMEM95*	001902-9913	SNV (missense)	Fleckvieh	[338]
Deficiency of uridine monophosphate synthase (DUMPS) (embryonic lethality)	AR	*UMPS*	000262-9913	SNV (nonsense)	Holstein	[339, 340]

Semen quality and male fertility	AR	*WDR19*		SNV (splice site)	Brown Swiss	[50]
Polled and multisystemic syndrome	AD*	*ZEB2*	001736-9913	gross deletion; 11bp deletion (frameshift)	Charolais, Fleckvieh	[341, 342]

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, MF = multifactorial, XR = X-linked recessive, OMIA = Online Mendelian Inheritance in Animals [65], ref = references

**de novo* mutation

[&]incomplete penetrance

Table A7: Newly identified candidate causative variants of the forward genetic approach.

						candidate variant [150, 155]			
disorder	**MOI**	**gene**	***gene name*** [343]	**chr**	**position**	**base change**	**amino acid change**	**transcript**	**effect**
Achondrogenesis, type II (Bulldog calf syndrome)	AD	*COL2A1*	*collagen type II alpha 1 chain*	5	32,308,734	c.3166G>A	p.Gly1056Ser	NM_001001135.3	missense, splice region
Achondrogenesis, type II (Bulldog calf syndrome)	AD	*COL2A1*	*collagen type II alpha 1 chain*	5	32,303,127	3513bp deletion	deletion of 10 exons		frameshift
Achondrogenesis, type II (Bulldog calf syndrome)	AD	*COL2A1*	*collagen type II alpha 1 chain*	5	32,301,911	6678bp deletion	deletion of 18 exons		frameshift
Achromatopsia	AR	*CNGB3*	*cyclic nucleotide gated channel subunit beta 3*	14	76,011,964	c.751G>A	p.Asp251Asn	XM_015474554.2	missense
Chondrodysplasia	AD	*FGFR3*	*fibroblast growth factor receptor 3*	6	116,767,863	c.2408G>T	p.Ter813Leuext*93	XM_024992994.1	stop lost
Classical Ehlers-Danlos syndrome, type II	AD	*COL5A2*	*collagen type V alpha 2 chain*	2	7,331,916	c.2366G>T	p.Gly789Val	XM_024979774	missense
Congenital cataract	AD	*ADAMTSL4*	*ADAMTS like 4*	3	20,146,737	c.2327G>A	p.Arg776His	NM_001101061.1	missense
Congenital neuromuscular channelopathy	AD	*KCNG1*	*potassium voltage-gated channel modifier subfamily G member 1*	13	78,918,850	c.1248G>T	p.Trp416Cys	NM_001205719.1	missense
Epidermolysis bullosa simplex	AD	*KRT5*	*keratin 5*	5	27,367,604	c.534_536delCAA	p.Asn178del	NM_001008663.1	in-frame deletion
Hemifacial microsomia	AR	*LAMB1*	*laminin subunit beta 1*	4	49,019,693	c.2002G>A	p.Arg668Cys	NM_001206519.1	missense

Hypohidrotic ectodermal dysplasia	XR	*EDA*	*ectodysplasin A*	X	80,382,423	52,780bp deletion	deletion of last 6 exons		knock-out
Ichthyosis congenita	AR	*FA2H*	*fatty acid 2-hydroxylase*	18	2,205,625	c.9dupC	p.Ala4ArgfsTer142	NM_001192455.1	frameshift
Syndromic ichthyosis	AR	*DSP*	*desmoplakin*	23	47,826,600	c.6893C>A	p.Ala2298Asp	NM_001192368.2	missense
Immunodeficiency with psoriasis-like skin alterations	AR	*IL17RA*	*interleukin 17 receptor A*	5	108,813,251	c.180delC	p.Cys61AlafsTer62	XM_015460734.2	frameshift
Osteogenesis imperfecta, type II	AD	*COL1A1*	*collagen type I alpha 1 chain*	19	36,473,359	c.3917T>A	p.Val1306Glu	NM_001034039.2	missense
Proportional dwarfism with facial dysplasia	AD	many		29		duplication of the entire chromosome			trisomy
Pulmonary hypoplasia and anasarca syndrome	AR	*ADAMTS3*	*ADAM metallopeptidase with thrombospondin type 1 motif 3*	6	87,462,016	c.1222C>T	p.His408Tyr	NM_001192797.1	missense
Pulmonary hypoplasia and anasarca syndrome	AD	many		20		duplication of the entire chromosome			trisomy
Skeletal-cardo-enteric dysplasia	AD	*MAP2K2*	*mitogen-activated protein kinase kinase 2*	7	19,923,991	c.535C>T	p.Arg179Trp	NM_001038071.2	missense

MOI = mode of inheritance, AR = autosomal recessive, AD = autosomal dominant, XR = X-linked recessive

Table A8: Identified haplotypes of the reverse genetic approach denoted Brown Swiss haplotypes (BH), Original Braunvieh haplotypes (OH), Holstein haplotypes (HH), and Simmental haplotypes (SH).

haplotype name	chr	start	end	dataset	AF (%)	proposed associated gene[a]	OMIA
BH1[b]	7	41,371,808	42,545,291	trio	3.02	*TCF3*	002513-9913
		41,376,851	42,569,144	pgp	3.66		
BH2[c]	19	9,726,237	10,819,756	trio	4.22	*TUBD1*	002513-9913
BH3	1	34,908,448	36,034,631	trio	2.05		002513-9913
		32,095,019	33,398,878	pgp	2.09		
BH4	1	76,587,512	77,861,430	trio	2.20	*GMNC*	002513-9913
		76,701,704	77,960,658	pgp			
BH5	2	57,787,148	58,768,161	trio	2.87	*LRP1B*	002513-9913
		55,350,431	56,566,658	pgp	2.84		
BH6	2	86,065,338	87,460,373	trio & pgp	3.41	*MARS2*	002513-9913
BH7	2	120,284,002	121,473,182	trio & pgp	1.93	*SPATA3*	002513-9913
BH8	2	127,933,873	129,004,898	trio	2.98		002513-9913
BH9	3	52,432,936	52,689,742	trio	1.43	*HFM1*	002513-9913
BH10	3	101,636,810	102,827,915	trio & pgp	2.26		002513-9913
BH11	4	108,811,998	109,953,666	trio	2.66		002513-9913
BH12	5	57,653,479	59,777,422	trio	2.32	*RNF41* and *LRP1*	002513-9913
		55,350,431	56,566,658	pgp	2.08		
BH13	5	110,617,654	111,664,596	trio	1.63		002513-9913
BH14	7	2,588,873	3,357,718	trio & pgp	3.04	*MRPL55*	002513-9913
BH15	7	33,954,389	34,956,540	trio	4.12		002513-9913
		33,761,341	34,703,029	pgp	2.23		
BH16	7	79,845,977	80,708,046	prio	2.20		002513-9913
BH17	11	82,034,654	82,881,173	trio	2.39	*PGGHG*	002513-9913
BH18	11	91,766,016	92,804,942	trio & pgp	2.54		002513-9913
BH19	11	103,891,372	105,003,404	trio	1.77		002513-9913
		103,910,192	105,006,092	pgp			
BH20	12	10,254,014	10,967,287	trio	2.02		002513-9913
BH21	12	74,954,939	75,817,277	trio	1.61		002513-9913
BH22	14	12,290,759	13,350,202	trio	2.64		002513-9913
BH23	17	69,786,512	70,645,226	trio	2.98		002513-9913
		70,347,783	71,461,810	pgp	2.53		
BH24	18	52,977,041	54,113,281	trio	2.32	*CPT1C*	002513-9913

BH25	19	35,808,547	36,844,178	trio	1.86	*SNF8*	002513-9913
BH26	19	47,874,708	48,724,082	trio	1.89		002513-9913
		44,273,561	45,438,908	pgp	1.51		
BH27	20	50,058,036	51,370,414	trio	2.56		002513-9913
		51,843,538	53,347,823	pgp	2.54		
BH28	21	19,368,258	20,362,870	trio	1.60		002513-9913
BH29	22	311,724	1,277,064	trio	2.07		002513-9913
BH30	23	39,600,366	40,609,442	trio & pgp	2.49		002513-9913
BH31	25	27,713,801	29,132,187	trio	1.80	*ITGAD* and *SPN*	002513-9913
		27,843,792	29,223,697	pgp	1.74		
BH32	26	9,215,188	10,180,467	trio & pgp	1.72		002513-9913
BH33	26	26,659,976	27,689,471	trio & pgp	2.72		002513-9913
BH34	26	31,353,340	32,429,589	trio	2.37	*ACSL5*	002513-9913
BH35	27	28,524,504	29,550,764	trio	1.80		002513-9913
BH36	28	26,661,955	27,743,768	trio	1.75		002513-9913
BH37	28	45,215,132	45,913,154	trio	2.07		002513-9913
BH38	29	14,497,752	15,554,643	trio	3.08		002513-9913
BH39	1	47,334,486	48,722,608	pgp	2.10	*IMPG2*	
BH40	1	60,571,934	61,511,334	pgp	1.75		
BH41	1	89,034,232	89,950,486	pgp	2.15		
BH42	2	27,617,503	28,482,635	pgp	1.80		
BH43	3	47,404,352	48,406,110	pgp	1.44		
BH44	4	1,281,131	2,168,698	pgp	1.59		
BH45	6	17,261,709	18,109,566	pgp	2.16		
BH46	6	112,761,523	113,596,244	pgp	2.97		
BH47	8	108,270,441	110,423,252	pgp	1.75		
BH48	9	6,723,284	7,824,441	pgp	1.61		
BH49	9	22,526,331	23,514,514	pgp	2.58		
BH50	9	89,606,305	90,789,128	pgp	1.71	*SYNE1*	
BH51	9	103,406,632	104,550,917	pgp	1.64		
BH52	10	21,853,744	25,057,987	pgp	2.65	*HCN4*	
BH53	10	97,990,595	98,716,261	pgp	1.51		
BH54	11	83,951,577	85,005,507	pgp	1.78		
BH55	12	23,595,402	24,816,308	pgp	2.85		
BH56	12	59,008,084	60,007,646	pgp	2.65		
BH57	12	68,078,290	68,922,617	pgp	2.46		
BH58	14	58,056,243	59,220,985	pgp	2.24		
BH59	18	61,266,875	62,181,523	pgp	2.75		
BH60	19	51,815,015	53,213,981	pgp	1.74		
BH61	20	22,655,940	23,695,033	pgp	1.92		
BH62	20	34,128,804	35,304,566	pgp	2.23		
BH63	20	42,147,837	43,447,120	pgp	1.84	*PDZD2*	
BH64	20	63,882,214	64,948,571	pgp	1.53		

BH65	22	18,499,781	19,588,204	pgp	2.82		
BH66	23	6,007,449	7,059,438	pgp	2.67		
BH67	23	24,736,986	25,781,822	pgp	1.74		
BH68	24	31,971,780	32,956,036	pgp	2.74		
BH69	25	35,189,206	36,281,889	pgp	2.15		
BH70	27	14,865,875	15,902,582	pgp	2.60		
BH71	29	16,595,399	17,452,624	pgp	3.09		
BH72	29	23,190,350	24,098,223	pgp	1.77		
HH18	1	103,019,622	104,375,908	pgp	2.39		
HH19[d]	1	139,873,827	140,642,688	pgp	3.00		
HH20	2	134,476,837	135,147,756	pgp	1.66		
HH21[e]	7	7,871,925	10,432,630	trio & pgp	5.46	*NOTCH3*	001908-9913
HH22	7	93,581,943	94,642,456	pgp	1.73		
HH23	8	15,561,949	16,862,751	pgp	2.20		
HH24	8	66,699,044	68,005,926	pgp	3.36		
HH3[f]	8	90,958,661	92,084,831	trio & pgp	1.38	*SMC2*	001824-9913
HH5[g]	9	90,928,417	91,823,561	trio & pgp	2.39	*TFB1M*	
HH25	10	86,876,435	87,772,797	trio	1.96	*RIOX1*	002547-9913
HH26	11	4,531,428	5,362,251	pgp	2.56		
HH27	13	7,083,081	8,296,886	pgp	1.82		
HH28	14	24,591,331	24,827,448	pgp	1.98		
HH29	14	58,128,474	59,238,055	pgp	3.14		
HH13[h]	18	60,931,980	62,100,899	trio	1.81	*KIR2DS1*	001836-9913
		62,070,448	63,044,863	pgp	2.06		
HH30	21	7,844,296	8,671,178	trio	2.13		
HH31	22	59,811,442	60,703,550	pgp	1.80		
HH32	23	32,773,236	33,726,454	trio	1.57		
		33,544,569	34,948,131	pgp	1.55		
HH33	24	44,692,769	45,677,988	pgp	2.82		
HH34	25	36,165,207	37,329,782	trio & pgp	1.86		
HH35	26	3,358,717	4,234,871	pgp	2.44	*PCDH15*	002548-9913
HH36	28	28,541,274	29,736,156	pgp	1.71		
		29,731,696	30,877,496	trio	2.25		
HH37	28	40,347,735	41,271,169	pgp	3.54		
HH38	29	14,242,732	15,316,961	trio	1.66		
OH2	2	1,005,580	1,614,673	trio	4.48	*TUBGCP5*	002514-9913
		1,029,782	1,647,935	pgp			
OH3	11	103,406,494	104,418,358	trio	5.76	*MYMK*	002514-9913
		104,201,626	105,231,051	pgp	6.50		
OH4	19	14,336,760	15,222,429	trio	3.87	*LIG3*	002514-9913
		14,354,974	15,240,920	pgp			
OH5	21	5,195,518	6,367,707	trio	4.56	*LYSMD4*	002514-9913
OH6	25	9,596,610	10,624,288	trio	3.87	*USP7*	002514-9913
OH7	6	4,963,003	6,358,918	trio	4.15	*BBS7*	
OH8	6	67,934,978	69,104,986	trio	4.46	*SCFD2*	
OH9	25	25,921,957	27,221,647	trio & pgp	3.52	*HSD3B7*	

SH1	1	65,792,796	66,890,930	trio & pgp	4.18		002501-9913
SH2	2	1,191,490	1,919,960	trio	4.00		002502-9913
SH3	7	104,784,181	105,639,721	pgp	4.14		002503-9913
SH4	11	29,389,110	30,352,564	pgp	4.67		002504-9913
SH5	12	46,828,195	47,836,535	trio	6.22	*DIS3*	002505-9913
		46,831,407	47,842,618	pgp			
SH6	14	11,304,160	12,372,654	pgp	3.85		002506-9913
SH7	16	6,120,019	8,011,283	trio	5.53		002507-9913
		6,121,692	8,060,022	pgp	10.60		
SH8	18	48,762,634	50,004,821	pgp	3.37	*CYP2B6*	002508-9913
		48,805,914	50,016,940	trio	3.38		
SH9	21	41,854,733	42,850,532	trio & pgp	4.33	*NUBPL*	002509-9913
SH10	24	41,944,725	43,140,179	pgp	3.24		002510-9913
SH11	28	38,936,504	40,068,646	trio	3.54		002511-9913

chr = chromosome, AF = allele frequency, OMIA = Online Mendelian Inheritance in Animals [65]
[a]according to NCBI Annotation Release 106
[b]previously described haplotype [103]
[c]previously described haplotype and associated gene [103, 119]
[d]haplotype co-localises with previously described haplotype HHB by Shuster *et al.* (1992) and Cole *et al.* (2018) [312, 344]
[e]haplotype described before as 175.5 and 07-126 by VanRaden et al. (2011) and Sahana et al. (2013) [103, 107]
[f]haplotype previously described by VanRaden et al. (2011), McClure et al. (2013), Sahana et al. (2013) and Wu et al. (2019) [103, 107, 108, 292]
[g]haplotype previously described by Schütz et al. (2016) and Fritz et al. (2013) [81, 124]
[h]haplotype previously described by Fritz et al. (2013) [109]

Table A9: Identified candidate causative variants of the reverse genetic approach denoted Brown Swiss haplotypes (BH), Original Braunvieh haplotypes (OH), Holstein haplotypes (HH), and Simmental haplotypes (SH).

haplotype		candidate variant [150, 155]						associated gene			
name	AF (%)	chr	position	base change	amino acid change	transcript	effect	gene	gene name [343]	associated disorder / function	OMIM / OMIA
BH1	3.0	7	43,997,146	c.208G>A	p.Asp70Asn	NM_001192698.1	missense variant	*TCF3*	*transcription factor 3*	Agammaglobulinemia 8; postnatal death	147141 / 001825-9913
BH2	4.2	19	10,833,921	c.629A>G	p.His210Arg	NM_001075470.2	missense variant	*TUBD1*	*tubulin delta 1*	embryonic lethal	607344 / 001939-9913
BH4[&]	2.2	1	76,381,812	c.733G>A	p.Gly245Arg	XM_002684863.5	missense variant	*GMNC*	*geminin coiled-coil domain containing*	infertility; postnatal lethality	614448
BH5[&]	2.9	2	56,650,799	c.4005C>T	p.Asn1335Asn	XM_024978797.1	synonymous variant	*LRP1B*	*LDL receptor related protein 1B*	embryonic lethality, preweaning lethality	608766
BH6	3.4	2	86,191,230	c.1553G>A	p.Arg518Gln	NM_001098971.1	missense variant	*MARS2*	*methionyl-tRNA synthetase 2, mitochondrial*	spastic ataxia 3, lethality	609728 / 002517-9913
BH7[&]	1.9	2	118,811,610	c.616C>T	p.Arg206Cys	NM_001076480.2	missense variant	*SPATA3*	*spermatogenesis associated 3*	sperm alterations; in vitro hypofertility	
BH9[&]	1.4	3	52,024,307	c.2341+4A>G		NM_001205576.2	splice region variant & intron variant	*HFM1*	*helicase for meiosis 1*	premature ovarian failure 9	615684
BH12[&]	2.3	5	57,129,352	c.852G>A	p.Gln284Gln	NM_001046525.1	synonymous variant	*RNF41*	*ring finger protein 41*	homozygous lethal	
BH12[&]	2.3	5	56,310,186	c.-68T>G		XM_024992496.1	5' UTR premature start codon gain variant	*LRP1*	*LDL receptor related protein 1*	Keratosis pilaris atrophicans, homozygous lethal	107770
BH14	3.0	7	2,996,436	c.169C>T	p.Arg57*	NM_001303490.1	stop gained	*MRPL55*	*mitochondrial ribosomal protein L55*	early pregnancy loss	611859 / 002518-9913
BH17[&]	2.4	11	82,957,204	c.41G>A	p.Arg14His	XM_010800733.3	missense variant	*PGGHG*	*protein-glucosylgalactosyl*	hydrolization of glucose; essential for catalytic actions	617032

									hydroxylysine glucosidase		
BH24	2.3	18	56,098,048	c.158G>A	p.Gly53Asp	XM_002695120.5	missense variant	*CPT1C*	*carnitine palmitoyltransferase 1C*	spastic paraplegia (lethal)	608846 / 002519-9913
BH25[&]	1.9	19	37,594,866	c.767C>T	p.Ala256Val	NM_001075211.1	missense variant	*SNF8*	*SNF8 subunit of ESCRT-II*	hirschsprungs disease; follicular development	610904
BH31[&]	1.8	25	27,455,769	c.239G>A	p.Arg80Gln	NM_001102496.2	missense variant	*ITGAD*	*integrin subunit alpha D*	abnormal immune system physiology	602453
BH31[&]	1.8	25	26,512,393	c.573T>G	p.Asp191Glu	NM_001103102.2	missense variant	*SPN*	*sialophorin*	spondylocostal dysostosis 5	182160
BH34	2.4	26	32,940,521	c.528C>G	p.Asn176Lys	NM_001075650.1	missense variant	*ACSL5*	*acyl-CoA synthetase long chain family member 5*	Lipid malabsorption	605677 / 002226-9913
BH39[&]	2.1	1	45,757,401	c.3106T>C	p.Tyr1036His	XM_024994100.1	missense variant	*IMPG2*	*interphotoreceptor matrix proteoglycan 2*	macular dystrophy, retinitis pigmentosa	607056 / 002289-9615
BH50[&]	1.7	9	89,319,503	c.16229C>T	p.Thr5410Ile	XM_024997066.1	missense variant	*SYNE1*	*spectrin repeat containing nuclear envelope protein 1*	Arthrogryposis multiplex congenita 3, Emery-Dreifuss muscular dystrophy, spinocerebellar ataxia; postnatal growth retardation	608441
BH52[&]	2.7	10	20,148,667	c.2167A>G	p.Arg723Gly	XM_024997938.1	missense variant	*HCN4*	*hyperpolarization activated cyclic nucleotide gated potassium channel 4*	Brugada syndrome 8, sick sinus syndrome 2; embryonic lethality in mice	605206
BH63[&]	1.8	20	41,522,590	c.3769G>A	p.Gly1257Arg	XM_024981517.1	missense variant	*PDZD2*	*PDZ domain containing 2*	cell adhesion; associated with prostate tumors	610697
HH3	1.4	8	93,753,358	c.3404T>C	p.Phe1135Ser	XM_015472668.2	missense variant	*SMC2*	*structural maintenance of chromosomes 2*	embryonic lethality	605576 / 001824-9913

HH13	1.8	18	62,758,881	c.475C>T	p.Gln159*	NM_001097567.1	stop gain variant	*KIR2DS1*	*killer cell immunoglobulin like receptor, two Ig domains and short cytoplasmic tail 1*	function in placentation and during pregnancy, integral component of membrane	604952 / 001836-9913
HH21	5.5	7	7,913,459	c.129_131delTTG	p.Cys44del	XM_003586246.3	disruptive in-frame deletion	*NOTCH3*	*notch receptor 3*	myofibrimatosis / cerebral arteriopathy with subcortical infarcts and leukoencephalopathy, type 1 (CADASIL1)	600276 / 001908-9913
HH25	2.0	10	84,938,370	c.396_425delGGCGCAGACCCCGGCGGCACGCTTGGTGGA	p.Ala133_Glu142del	NM_001099702.1	disruptive in-frame deletion	*RIOX1*	*ribosomal oxygenase 1*	chromatin organization, negative regulation of osteoblast differentiation, negative regulation of transcription, DNA-templated	611919 / 002547-9913
HH35	2.4	26	5,325,675	c.2599C>G	p.Leu867Val	XM_015460562.2	missense variant	*PCDH15*	*protocadherin related 15*	Usher syndrome, deafness	605514 / 002548-9913
OH2	4.5	2	1,268,426	c.311C>A	p.Thr104Lys	NM_001102495.1	missense variant & splice region variant	*TUBGCP5*	*tubulin gamma complex associated protein 5*	important role in the formation of the mitotic spindles	608147 / 002515-9913
OH3[&]	5.8	11	104,408,774	c.-762G>T		XM_024999145.1	5' UTR premature start codon gain variant	*MYMK*	*myomaker, myoblast fusion factor*	Carey-Fineman-Ziter syndrome; congenital myopathy	615345
OH4	3.9	19	15,080,335	c.2483_2484+4delAGGTGC	p.Lys828fs	NM_001038107.2	frameshift variant & splice donor variant & intron variant	*LIG3*	*DNA ligase 3*	embryonic lethality	600940 / 002516-9913

OH5[&]	4.6	21	6,825,663	c.460C>G	p.His154Asp	NM_001081528.1	missense variant	*LYSMD4*	*LysM domain containing 4*	coronary artery calcification	
OH6[&]	3.9	25	7,724,151	c.2500G>C	p.Glu834Gln	XM_005224667.4	missense variant	*USP7*	*ubiquitin specific peptidase 7*	Hao-Fountain syndrome	602519
OH7[&]	4.2	6	3,393,947	c.377G>A	p.Ser126Asn	NM_001191346.2	missense variant	*BBS7*	*Bardet-Biedl syndrome 7*	Bardet-Biedl syndrome 7; ciliopathic multisyndromic disorder	607590 / 002216-9544
OH8[&]	4.5	6	68,551,351	c.1904C>T	p.Thr635Ile	NM_001078005.1	missense variant	*SCFD2*	*sec1 family domain containing 2*	preweaning lethality	
OH9[&]	3.5	25	27,119,330	c.259C>T	p.Arg87Cys	XM_015460381.2	missense variant	*HSD3B7*	*hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7*	Bile acid synthesis defect	607764
SH5	6.2	12	47,511,687	c.2032dupA	p.Ile678AsnTer2	XM_025000110.1	frameshift variant	*DIS3*	*DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease*	RNA processing; preweaning lethality; Perlman syndrome	607533 / 002505-9913
SH8	3.4	18	50,296,371	c.938T>A	p.Ile313Asn	NM_001075173.1	missense variant	*CYP2B6*	*cytochrome P450 family 2 subfamily B member 6*	poor metabolism of efavirenz, central nervous system toxicity	123930 / 002508-9913
SH9	4.3	21	42,154,344	c.428C>A	p.Ser143Tyr	NM_001193042.1	missense variant	*NUBPL*	*nucleotide binding protein like*	Mitochondrial complex I deficiency, nuclear type 21; prenatal lethality prior to heart atrial spation; preweaning lethality, complete penetrance	613621 / 002509-9913

AF = allele frequency, OMIM = Online Mendelian Inheritance in Man [345], OMIA = Online Mendelian Inheritance in Animals [65],
[&]low confidence variants

9 References

1. Pitt D, Sevane N, Nicolazzi EL, MacHugh DE, Park SDE, Colli L, Martinez R, Bruford MW, Orozco-terWengel P (2019) Domestication of cattle: Two or three events? Evol Appl 12:123–136
2. Kothieringer K, Walser C, Dietre B, Reitmaier T, Haas JN, Lambers K (2015) High impact: early pastoralism and environmental change during the Neolithic and Bronze Age in the Silvretta Alps (Switzerland/Austria) as evidenced by archaeological, palaeoecological and pedological proxies. Zeitschrift für Geomorphol Suppl Issues 59:177–198
3. Robinson TP, Wint GRW, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, Cinardi G, D'Aietti L, Hay SI, Gilbert M (2014) Mapping the global distribution of livestock. PLoS One 9:e96084
4. Braunvieh Schweiz (2021) Pictures of Brown Swiss and Original Braunvieh cows. received from Martin Rust directly
5. Mutterkuh Schweiz (2021) Picture of a Limousin cow. received from Svenja Strasser directly
6. Verein Original Simmentaler (2021) Homepage Simmental Original. https://www.simmentaler-original.ch/de/die-rasse/#&gid=1&pid=1. Accessed 29 Oct 2021
7. Holstein Switzerland (2021) Homepage Holstein Switzerland. https://www.holstein.ch/de/die-besten-leistungen-august-2021/. Accessed 29 Oct 2021
8. Fleckvieh IG Swiss (2021) Homepage IG Swiss Fleckvieh. https://www.swissfleckvieh.ch. Accessed 29 Oct 2021
9. Identitas AG (2021) identitas. In: Tierstatistik. https://www.identitas.ch. Accessed 27 Oct 2021
10. Hagger C (2005) Estimates of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. J Anim Breed Genet 122:405–413
11. Signer-Hasler H, Burren A, Neuditschko M, Frischknecht M, Garrick D, Stricker C, Gredler B, Bapst B, Flury C (2017) Population structure and genomic inbreeding in nine Swiss dairy cattle populations. Genet Sel Evol 49:83
12. Stranzinger GF, Steiger D, Kneubühler J, Hagger C (2007) Y chromosome polymorphism in various breeds of cattle (Bos taurus) in Switzerland. J Appl Genet 48:241–245
13. Holstein Switzerland (2020) Geschäftsbericht 2020. https://www.holstein.ch/wp-content/uploads/rapport_de_gestion2020_Online.pdf. Accessed 27 Oct 2021
14. Swissherdbook (2020) Herdebuch- und Mitgliederbestand. https://www.swissherdbook.ch/fileadmin/Domain1/PDF_Dokumente/05-Statistiken-Formulare/53-

Jahresstatistik/1_Wichtigste_Zahlen/D_sh_JS_2020_HBZahlen_web.pdf. Accessed 27 Oct 2021

15. Braunvieh Schweiz (2020) Geschäftsbericht 2020. https://homepage.braunvieh.ch/wp-content/uploads/2021/06/GB-Braunvieh-Schweiz-2020-D.pdf. Accessed 27 Oct 2021
16. Mutterkuh Schweiz (2020) Mutterkuh Schweiz Herdebuch. https://www.mutterkuh.ch/de/herdebuch/rasse/limousin. Accessed 27 Oct 2021
17. Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Essex, 4th ed. Harlow
18. Goddard ME, Hayes BJ, Meuwissen THE (2010) Genomic selection in livestock populations. Genet Res (Camb) 92:413–421
19. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. J Dairy Sci 92:433–443
20. Pryce JE, Daetwyler HD (2011) Designing dairy cattle breeding schemes under genomic selection: a review of international research. Anim Prod Sci 52:107–114
21. Schefers JM, Weigel KA (2012) Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. Anim Front 2:4–9
22. Wiggans GR, Cole JB, Hubbard SM, Sonstegard TS (2017) Genomic selection in dairy cattle: the USDA experience. Annu Rev Anim Biosci 5:309–327
23. Lillehammer M, Meuwissen THE, Sonesson AK (2011) Genomic selection for maternal traits in pigs. J Anim Sci 89:3908–3916
24. Samorè AB, Fontanesi L (2016) Genomic selection in pigs: state of the art and perspectives. Ital J Anim Sci 15:211–232
25. Wolc A (2015) Genomic selection in layer and broiler breeding. Lohmann Inf 49:1–11
26. Fulton JE (2012) Genomic selection for poultry breeding. Anim Front 2:30–36
27. Zenger KR, Khatkar MS, Jones DB, Khalilisamani N, Jerry DR, Raadsma HW (2019) Genomic selection in aquaculture: application, limitations and opportunities with special reference to marine shrimp and pearl oysters. Front Genet 9:693
28. Sonesson AK, Meuwissen THE (2009) Testing strategies for genomic selection in aquaculture breeding programs. Genet Sel Evol 41:1–9
29. Robertson A (1954) Inbreeding and performance in British Friesian cattle. In: Proc. Br. Soc. Anim. Prod. pp 87–92
30. Biffani S, Samoré AB, Canavesi F (2002) Inbreeding depression for production, reproduction and functional traits in Italian Holstein cattle. Proc 7th world Congr Genet Appl to Livest Prod 19–23
31. Doekes HP, Veerkamp RF, Bijma P, de Jong G, Hiemstra SJ, Windig JJ

(2019) Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein–Friesian dairy cattle. Genet Sel Evol 51:54

32. Doekes HP, Bijma P, Windig JJ (2021) How depressing is inbreeding? A meta-analysis of 30 years of research on the effects of inbreeding in livestock. Genes (Basel) 12:926

33. Burrow HM (1993) The effects of inbreeding in beef cattle. In: Anim. Breed. Abstr. pp 737–751

34. Senczuk G, Mastrangelo S, Ciani E, Battaglini L, Cendron F, Ciampolini R, Crepaldi P, Mantovani R, Bongioni G, Pagnacco G, Portolano B, Rossoni A, Pilla F, Cassandro M (2020) The genetic heritage of Alpine local cattle breeds using genomic SNP data. Genet Sel Evol 52:40

35. Bhati M, Kadri NK, Crysnanto D, Pausch H (2020) Assessing genomic diversity and signatures of selection in Original Braunvieh cattle using whole-genome sequencing data. BMC Genomics 21:27

36. Thibier M, Wagner H-G (2002) World statistics for artificial insemination in cattle. Livest Prod Sci 74:203–212

37. Berry DP, Ring SC, Twomey AJ, Evans RD (2020) Choice of artificial insemination beef bulls used to mate with female dairy cattle. J Dairy Sci 103:1701–1710

38. Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF (2017) A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. J Dairy Sci 100:10251–10271

39. Hoffman JM, Valencak TG (2020) A short life on the farm: aging and longevity in agricultural, large-bodied mammals. GeroScience 42:909–922

40. Pryce JE, Woolaston R, Berry DP, Wall E, Winters M, Butler R, Shaffer M (2014) World trends in dairy cow fertility. In: World Congr. Genet. Appl. to Livest. Prod. Vancouver, BC, Canada, p 6

41. Walsh SW, Williams EJ, Evans ACO (2011) A review of the causes of poor fertility in high milk producing dairy cows. Anim Reprod Sci 123:127–138

42. Pryce JE, Royal MD, Garnsworthy PC, Mao IL (2004) Fertility in the high-producing dairy cow. Livest Prod Sci 86:125–135

43. Ma L, Sonstegard TS, Cole JB, Vantassell CP, Wiggans GR, Crooker BA, Tan C, Prakapenka D, Liu GE, Da Y (2019) Genome changes due to artificial selection in U.S. Holstein cattle. BMC Genomics 20:1–14

44. Zobel R, Tkalčić S, Pipal I, Buić V (2011) Incidence and factors associated with early pregnancy losses in Simmental dairy cows. Anim Reprod Sci 127:121–125

45. Seegers H, Beaudeau F, Fourichon C, Bareille N (1998) Reasons for culling in French Holstein cows. Prev Vet Med 36:257–271

46. Rilanto T, Reimus K, Orro T, Emanuelson U, Viltrop A, Mõtus K (2020) Culling reasons and risk factors in Estonian dairy cows. BMC Vet Res 16:173

47. Barth AD (1993) Factors affecting fertility with artificial insemination. Vet

Clin North Am Food Anim Pract 9:275–289

48. Demetrio DGB, Santos RM, Demetrio CGB, Vasconcelos JLM (2007) Factors affecting conception rates following artificial insemination or embryo transfer in lactating holstein cows. J Dairy Sci 90:5073–5082

49. Vanroose G, de Kruif A, Van Soom A (2000) Embryonic mortality and embryo–pathogen interactions. Anim Reprod Sci 60–61:131–143

50. Hiltpold M, Niu G, Kadri NK, Crysnanto D, Fang Z-H, Spengeler M, Schmitz-Hsu F, Fuerst C, Schwarzenbacher H, Seefried FR, Seehusen F, Witschi U, Schnieke A, Fries R, Bollwein H, Flisikowski K, Pausch H (2020) Activation of cryptic splicing in bovine WDR19 is associated with reduced semen quality and male fertility. PLOS Genet 16:e1008804

51. Hiltpold M, Kadri NK, Janett F, Witschi U, Schmitz-Hsu F, Pausch H (2021) Autosomal recessive loci contribute significantly to quantitative variation of male fertility in a dairy cattle population. BMC Genomics 22:225

52. Hiltpold M, Janett F, Mapel XM, Kadri NK, Fang Z-H, Schwarzenbacher H, Seefried F, Spengeler M, Witschi U, Pausch H (2022) A 1-bp deletion in bovine QRICH2 causes low sperm count and immotile sperm with multiple morphological abnormalities. Genet Sel Evol 54:1–14

53. Nani JP, Rezende FM, Peñagaricano F (2019) Predicting male fertility in dairy cattle using markers with large effect and functional annotation data. BMC Genomics 20:258

54. Nicolini P, Amorín R, Han Y, Peñagaricano F (2018) Whole-genome scan reveals significant non-additive effects for sire conception rate in Holstein cattle. BMC Genet 19:14

55. Abdollahi-Arpanahi R, Pacheco HA, Peñagaricano F (2021) Targeted sequencing reveals candidate causal variants for dairy bull subfertility. Anim. Genet.

56. Fortes MRS, DeAtley KL, Lehnert SA, Burns BM, Reverter A, Hawken RJ, Boe-Hansen G, Moore SS, Thomas MG (2013) Genomic regions associated with fertility traits in male and female cattle: Advances from microsatellites to high-density chips and beyond. Anim Reprod Sci 141:1–19

57. Widmer S, Seefried FR, von Rohr P, Häfliger IM, Spengeler M, Drögemüller C (2021) A major QTL at the LHCGR/FSHR locus for multiple birth in Holstein cattle. Genet Sel Evol 53:1–15

58. Schmutz SM, Moker JS, Clark EG, Orr JP (1996) Chromosomal aneuploidy associated with spontaneous abortions and neonatal losses in cattle. J Vet Diagnostic Investig 8:91–95

59. Hu ZL, Park CA, Reecy JM (2019) Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Res 47:D701–D710

60. Frischknecht M, Bapst B, Seefried FR, Signer-Hasler H, Garrick D, Stricker C, Intergenomics C, Fries R, Russ I, Solkner J, Bieber A, Strillacci MG, Gredler-Grandl B, Flury C (2017) Genome-wide association studies of fertility and calving traits in Brown Swiss cattle using imputed whole-

genome sequences. BMC Genomics 18:910

61. Mock T, Mee JF, Dettwiler M, Rodriguez-Campos S, Hüsler J, Michel B, Häfliger IM, Drögemüller C, Bodmer M, Hirsbrunner G (2020) Evaluation of an investigative model in dairy herds with high calf perinatal mortality rates in Switzerland. Theriogenology 148:48–59

62. Mohr OL, Wriedt C (1928) Hairless, a new recessive lethal in cattle. J Genet 19:315–336

63. Wriedt C, Mohr OL (1928) Amputated, a recessive lethal in cattle; with a discussion on the bearing of lethal factors on the principles of live stock breeding. J Genet 20:187–215

64. Watson JD, Crick FHC (1953) The structure of DNA. In: Cold Spring Harb. Symp. Quant. Biol. pp 123–131

65. OMIA (2021) Online Mendelian Inheritance in Animals. https://omia.org/home/. Accessed 31 Oct 2021

66. Georges M, Charlier C, Hayes B (2019) Harnessing genomic information for livestock improvement. Nat Rev Genet 20:135–156

67. Alsaaod M, Locher I, Jores J, Grimm P, Brodard I, Steiner A, Kuhnert P (2019) Detection of specific Treponema species and Dichelobacter nodosus from digital dermatitis (Mortellaro's disease) lesions in Swiss cattle. Schweiz Arch Tierheilkd 161:207–215

68. Bianchi M V, Silveira S, Mosena ACS, de Souza SO, Konradt G, Canal CW, Driemeier D, Pavarini SP (2019) Pathological and virological features of skin lesions caused by BVDV in cattle. Braz J Microbiol 50:271–277

69. Wong R, Geyer S, Weninger W, Guimberteau J, Wong JK (2016) The dynamic anatomy and patterning of skin. Exp Dermatol 25:92–98

70. Dittmer KE, Thompson KG (2015) Approach to investigating congenital skeletal abnormalities in livestock. Vet Pathol 52:851–861

71. Cavanagh JAL, Tammen I, Windsor PA, Bateman JF, Savarirayan R, Nicholas FW, Raadsma HW (2007) Bulldog dwarfism in Dexter cattle is caused by mutations in ACAN. Mamm Genome 18:808–814

72. Windsor PA, Kessell AE, Finnie JW (2011) Neurological diseases of ruminant livestock in Australia. V: congenital neurogenetic disorders of cattle. Aust Vet J 89:394–401

73. Mömke S, Distl O (2007) Bilateral convergent strabismus with exophthalmus (BCSE) in cattle: An overview of clinical signs and genetic traits. Vet J 173:272–277

74. Murgiano L, Jagannathan V, Calderoni V, Joechler M, Gentile A, Drögemüller C (2014) Looking the cow in the eye: Deletion in the NID1 gene is associated with recessive inherited cataract in Romagnola cattle. PLoS One 9:e110628

75. Hollmann AK, Dammann I, Wemheuer WM, Wemheuer WE, Chilla A, Tipold A, Schulz-Schaeffer WJ, Beck J, Schütz E, Brenig B (2017) Morgagnian cataract resulting from a naturally occurring nonsense mutation elucidates a role of CPAMD8 in mammalian lens development.

PLoS One 12:e0180665–e0180665

76. Michot P, Chahory S, Marete A, Grohs C, Dagios D, Donzel E, Aboukadiri A, Deloche M-C, Allais-Bonnet A, Chambrial M, Barbey S, Genestout L, Boussaha M, Danchin-Burge C, Fritz S, Boichard D, Capitan A (2016) A reverse genetic approach identifies an ancestral frameshift mutation in RP1 causing recessive progressive retinal degeneration in European cattle breeds. Genet Sel Evol 48:56

77. Schmutz SM, Berryere TG, Ciobanu DC, Mileham AJ, Schmidtz BH, Fredholm M (2004) A form of albinism in cattle is caused by a tyrosinase frameshift mutation. Mamm Genome 15:62–67

78. Rothammer S, Kunz E, Seichter D, Krebs S, Wassertheurer M, Fries R, Brem G, Medugorac I (2017) Detection of two non-synonymous SNPs in SLC45A2 on BTA20 as candidate causal mutations for oculocutaneous albinism in Braunvieh cattle. Genet Sel Evol 49:1–8

79. Kipp S, Segelke D, Schierenbeck S, Reinhardt F, Reents R, Wurmser C, Pausch H, Fries R, Thaller G, Tetens J, Pott J, Haas D, Raddatz BB, Hewicker-Trautwein M, Proios I, Schmicke M, Grünberg W (2016) Identification of a haplotype associated with cholesterol deficiency and increased juvenile mortality in Holstein cattle. J Dairy Sci 99:8915–8931

80. Menzi F, Besuchet-Schmutz N, Fragnière M, Hofstetter S, Jagannathan V, Mock T, Raemy A, Studer E, Mehinagic K, Regenscheit N, Meylan M, Schmitz-Hsu F, Drögemüller C (2016) A transposable element insertion in APOB causes cholesterol deficiency in Holstein cattle. Anim Genet 47:253–257

81. Schütz E, Wehrhahn C, Wanjek M, Bortfeld R, Wemheuer WE, Beck J, Brenig B (2016) The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. PLoS One 11:e0154602

82. Artesi M, Hahaut V, Cole B, Lambrechts L, Ashrafi F, Marçais A, Hermine O, Griebel P, Arsic N, van der Meer F (2019) Pooled CRISPR Inverse PCR sequencing (PCIP-seq): simultaneous sequencing of retroviral insertion points and the integrated provirus with long reads. bioRxiv 558130

83. Teare MD, Santibañez Koref MF (2014) Linkage analysis and the study of Mendelian disease in the era of whole exome and genome sequencing. Brief Funct Genomics 13:378–383

84. Ott J, Wang J, Leal SM (2015) Genetic linkage analysis in the age of whole-genome sequencing. Nat Rev Genet 16:275–284

85. Vahidnezhad H, Youssefian L, Jazayeri A, Uitto J (2018) Research techniques made simple: Genome-wide homozygosity/autozygosity mapping Is a powerful tool for identifying candidate genes in autosomal recessive genetic diseases. J Invest Dermatol 138:1893–1900

86. Alfares A, Alsubaie L, Aloraini T, Alaskar A, Althagafi A, Alahmad A, Rashid M, Alswaid A, Alothaim A, Eyaid W (2020) What is the right sequencing approach? Solo VS extended family analysis in consanguineous

populations. BMC Med Genomics 13:1–8

87. Goldmann JM, Veltman JA, Gilissen C (2019) De novo mutations reflect development and aging of the human germline. Trends Genet 35:828–839

88. Biesecker LG, Spinner NB (2013) A genomic view of mosaicism and human disease. Nat Rev Genet 14:307–320

89. Bourneuf E, Otz P, Pausch H, Jagannathan V, Michot P, Grohs C, Piton G, Ammermuller S, Deloche MC, Fritz S, Leclerc H, Pechoux C, Boukadiri A, Hoze C, Saintilan R, Crechet F, Mosca M, Segelke D, Guillaume F, Bouet S, Baur A, Vasilescu A, Genestout L, Thomas A, Allais-Bonnet A, Rocha D, Colle MA, Klopp C, Esquerre D, Wurmser C, Flisikowski K, Schwarzenbacher H, Burgstaller J, Brugmann M, Dietschi E, Rudolph N, Freick M, Barbey S, Fayolle G, Danchin-Burge C, Schibler L, Bed'Hom B, Hayes BJ, Daetwyler HD, Fries R, Boichard D, Pin D, Drogemuller C, Capitan A (2017) Rapid discovery of de novo deleterious mutations in cattle enhances the value of livestock as model species. Sci Rep 7:1–19

90. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, Genomes Project Analysis G (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158

91. Häfliger IM (2021) Program: filterVariants. https://github.com/ihaefliger/program_filterVariants. Accessed 7 Sep 2021

92. Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nat Rev Genet 7:85–97

93. Bickhart DM, Hou Y, Schroeder SG, Alkan C, Cardone MF, Matukumalli LK, Song J, Schnabel RD, Ventura M, Taylor JF, Garcia JF, Van Tassell CP, Sonstegard TS, Eichler EE, Liu GE (2012) Copy number variation of individual cattle genomes using next-generation sequencing. Genome Res 22:778–790

94. Boussaha M, Esquerre D, Barbieri J, Djari A, Pinton A, Letaief R, Salin G, Escudie F, Roulet A, Fritz S, Samson F, Grohs C, Bernard M, Klopp C, Boichard D, Rocha D (2015) Genome-wide study of structural variants in bovine Holstein, Montbeliarde and Normande dairy breeds. PLoS One 10:e0135931

95. Keel BN, Keele JW, Snelling WM (2017) Genome-wide copy number variation in the bovine genome detected using low coverage sequence of popular beef breeds. Anim Genet 48:141–150

96. Mesbah-Uddin M, Guldbrandtsen B, Iso-Touru T, Vilkki J, De Koning DJ, Boichard D, Lund MS, Sahana G (2017) Genome-wide mapping of large deletions and their population-genetic properties in dairy cattle. DNA Res 25:49–59

97. Butty AM, Chud TCS, Miglior F, Schenkel FS, Kommadath A, Krivushin K, Grant JR, Häfliger IM, Drögemüller C, Cánovas A (2020) High confidence copy number variants identified in Holstein dairy cattle from whole genome sequence and genotype array data. Sci Rep 10:1–13

98. Butty AM, Chud TCS, Cardoso DF, Lopes LSF, Miglior F, Schenkel FS,

Cánovas A, Häfliger IM, Drögemüller C, Stothard P (2021) Genome-wide association study between copy number variants and hoof health traits in Holstein dairy cattle. J Dairy Sci 104:8050–8061

99. Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y (2019) Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol 20:117

100. Häfliger IM (2019) Program: coveragePlot_region. https://github.com/ihaefliger/program_coveragePlot_region. Accessed 6 Aug 2019

101. Grahofer A, Letko A, Häfliger IM, Jagannathan V, Ducos A, Richard O, Peter V, Nathues H, Drögemüller C (2019) Chromosomal imbalance in pigs showing a syndromic form of cleft palate. BMC Genomics 20:1–11

102. Simon R, Lischer HEL, Pieńkowska-Schelling A, Keller I, Häfliger IM, Letko A, Schelling C, Lühken G, Drögemüller C (2020) New genomic features of the polled intersex syndrome variant in goats unraveled by long-read whole-genome sequencing. Anim Genet 51:439–448

103. VanRaden PM, Olson KM, Null DJ, Hutchison JL (2011) Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J Dairy Sci 94:6153–6161

104. Hoff JL, Decker JE, Schnabel RD, Taylor JF (2017) Candidate lethal haplotypes and causal mutations in Angus cattle. BMC Genomics 18:799

105. Casellas J, Cañas-Alvarez JJ, González-Rodríguez A, Puig-Oliveras A, Fina M, Piedrafita J, Molina A, Diaz C, Baro JA, Varona L (2017) Bayesian analysis of parent-specific transmission ratio distortion in seven Spanish beef cattle breeds. Anim Genet 48:93–96

106. Jenko J, McClure MC, Matthews D, McClure J, Johnsson M, Gorjanc G, Hickey JM (2019) Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal and semi-lethal alleles with pleiotropic effects on economically important traits in beef cattle. Genet Sel Evol 51:9

107. Sahana G, Nielsen US, Aamand GP, Lund MS, Guldbrandtsen B (2013) Novel harmful recessive haplotypes identified for fertility traits in Nordic Holstein cattle. PLoS One 8:8–12

108. Wu X, Mesbah-Uddin M, Guldbrandtsen B, Lund MS, Sahana G (2019) Haplotypes responsible for early embryonic lethality detected in Nordic Holsteins. J Dairy Sci 102:11116–11123

109. Fritz S, Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, Grohs C, Weiss B, Boussaha M, Esquerre D, Klopp C, Rocha D, Boichard D (2013) Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. PLoS One 8:e65550

110. Pausch H, Schwarzenbacher H, Burgstaller J, Flisikowski K, Wurmser C, Jansen S, Jung S, Schnieke A, Wittek T, Fries R (2015) Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. BMC Genomics 16:312

111. Cooper TA, Wiggans GR, Null DJ, Hutchison JL, Cole JB (2014) Genomic

evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. J Dairy Sci 97:3878–3882

112. Derks MFL, Megens HJ, Bosse M, Lopes MS, Harlizius B, Groenen MAM (2017) A systematic survey to identify lethal recessive variation in highly managed pig populations. BMC Genomics 18:858

113. Haggman J, Uimari P (2017) Novel harmful recessive haplotypes for reproductive traits in pigs. J Anim Breed Genet 134:129–135

114. Howard DM, Pong-Wong R, Knap PW, Woolliams JA (2017) Use of haplotypes to identify regions harbouring lethal recessive variants in pigs. Genet Sel Evol 49:1–10

115. Derks MFL, Megens HJ, Bosse M, Visscher J, Peeters K, Bink M, Vereijken A, Gross C, de Ridder D, Reinders MJT, Groenen MAM (2018) A survey of functional genomic variation in domesticated chickens. Genet Sel Evol 50:17

116. Todd ET, Thomson PC, Hamilton NA, Ang RA, Lindgren G, Viklund Å, Eriksson S, Mikko S, Strand E, Velie BD (2020) A genome-wide scan for candidate lethal variants in Thoroughbred horses. Sci Rep 10:13153

117. Venhoranta H, Pausch H, Flisikowski K, Wurmser C, Taponen J, Rautala H, Kind A, Schnieke A, Fries R, Lohi H, Andersson M (2014) In frame exon skipping in UBE3B is associated with developmental disorders and increased mortality in cattle. BMC Genomics 15:890

118. Null DJ, Huthcins JL, Bickhart DM, VanRaden PM, Cole JB (2017) Discovery of a haplotype affecting fertility in Ayrshire dairy cattle and identification of a putative causal variant. J Dairy Sci 100:199

119. Schwarzenbacher H, Burgstaller J, Seefried FR, Wurmser C, Hilbe M, Jung S, Fuerst C, Dinhopl N, Weissenböck H, Fuerst-Waltl B, Dolezal M, Winkler R, Grueter O, Bleul U, Wittek T, Fries R, Pausch H (2016) A missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and Fleckvieh cattle. BMC Genomics 17:400

120. Joller S, Stettler M, Locher I, Dettwiler M, Seefried F, Meylan M, Drögemüller C (2018) Fanconi-bickel-syndrom: A novel genetic disease in Original Braunvieh. Schweiz Arch Tierheilkd 160:179–184

121. Adams HA, Sonstegard TS, VanRaden PM, Null DJ, Van Tassell CP, Larkin DM, Lewin HA (2016) Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in Holstein dairy cattle. J Dairy Sci 99:6693–6701

122. Ortega MS, Bickhart DM, Clark KN, Null DJ, Hutchison JL, McClure JC, Cole JB (2021) Truncation of IFT80 causes early embryonic loss in cattle. bioRxiv 2021.07.02.450952

123. McClure MC, Bickhart D, Null D, VanRaden P, Xu LY, Wiggans G, Liu G, Schroeder S, Glasscock J, Armstrong J, Cole JB, Van Tassell CP, Sonstegard TS (2014) Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2, and HH3 reveal a putative Causative mutation in SMC2 for HH3. PLoS One 9:e92769

124. Fritz S, Hoze C, Rebours E, Barbat A, Bizard M, Chamberlain A,

Escouflaire C, Vander Jagt C, Boussaha M, Grohs C, Allais-Bonnet A, Philippe M, Vallée A, Amigues Y, Hayes BJ, Boichard D, Capitan A (2018) An initiator codon mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. J Dairy Sci 101:6220–6231

125. Hozé C, Escouflaire C, Mesbah-Uddin M, Barbat A, Boussaha M, Deloche MC, Boichard D, Fritz S, Capitan A (2020) A splice site mutation in CENPU is associated with recessive embryonic lethality in Holstein cattle. J Dairy Sci 103:607–612

126. Sonstegard TS, Cole JB, VanRaden PM, Van Tassell CP, Null DJ, Schroeder SG, Bickhart D, McClure MC (2013) Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. PLoS One 8:e54872

127. Michot P, Fritz S, Barbat A, Boussaha M, Deloche MC, Grohs C, Hoze C, Le Berre L, Le Bourhis D, Desnoes O, Salvetti P, Schibler L, Boichard D, Capitan A (2017) A missense mutation in PFAS (phosphoribosylformylglycinamidine synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbeliarde dairy cattle. J Dairy Sci 100:8176–8187

128. Mesbah-Uddin M, Hoze C, Michot P, Barbat A, Lefebvre R, Boussaha M, Sahana G, Fritz S, Boichard D, Capitan A (2019) A missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in French Normande cattle. J Dairy Sci 102:6340–6356

129. Charlier C, Li W, Harland C, Littlejohn M, Coppieters W, Creagh F, Davis S, Druet T, Faux P, Guillaume F, Karim L, Keehan M, Kadri NK, Tamma N, Spelman R, Georges M (2016) NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. Genome Res 26:1333–1341

130. Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 76:887–893

131. Derks MFL, Gjuvsland AB, Bosse M, Lopes MS, van Son M, Harlizius B, Tan BF, Hamland H, Grindflek E, Groenen MAM, Megens HJ (2019) Loss of function mutations in essential genes cause embryonic lethality in pigs. PLoS Genet 15:e1008055

132. Derks MFL, Lopes MS, Bosse M, Madsen O, Dibbits B, Harlizius B, Groenen MAM, Megens HJ (2018) Balancing selection on a recessive lethal deletion with pleiotropic effects on two neighboring genes in the porcine genome. PLoS Genet 14:e1007661

133. Subramanian S (2021) Deleterious protein-coding variants in diverse cattle breeds of the world. Genet Sel Evol 53:80

134. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

135. Garrick DJ, Taylor JF, Fernando RL (2009) Deregressing estimated breeding values and weighting information for genomic regression

analyses. Genet Sel Evol 41:55

136. Teo YY (2008) Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. Curr Opin Lipidol 19:133–143

137. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188

138. Bland JM, Altman DG (1995) Multiple significance tests: the Bonferroni method. Bmj 310:170

139. Turner S (2018) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J Open Source Softw 34:1–2

140. R Core Team (2018) R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria

141. EMBL-EBI (2019) European Nucleotide Archive. https://www.ebi.ac.uk. Accessed 25 May 2019

142. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J (2013) From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinforma 43:11.10.1-11.10.33

143. Hayes BJ, Daetwyler HD (2019) 1000 Bull Genomes Project to map simple and complex genetic traits in cattle: Applications and outcomes. Annu Rev Anim Biosci 7:89–102

144. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brondum RF, Liao X, Djari A, Rodriguez SC, Grohs C, Esquerre D, Bouchez O, Rossignol MN, Klopp C, Rocha D, Fritz S, Eggen A, Bowman PJ, Coote D, Chamberlain AJ, Anderson C, VanTassell CP, Hulsegge I, Goddard ME, Guldbrandtsen B, Lund MS, Veerkamp RF, Boichard DA, Fries R, Hayes BJ (2014) Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat Genet 46:858–865

145. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890

146. Babraham Bioinformatics (2016) FastQC. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 16 Oct 2017

147. Ewels P, Magnusson M, Lundin S, Käller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048

148. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760

149. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, Rowan TN, Low WY, Zimin A, Couldrey C, Hall R, Li W, Rhie A, Ghurye J, McKay SD, Thibaud-Nissen F, Hoffman J, Murdoch BM, Snelling WM, McDaneld TG, Hammond JA, Schwartz JC, Nandolo W, Hagen DE, Dreischer C, Schultheiss SJ, Schroeder SG, Phillippy AM, Cole JB, Van Tassell CP, Liu

G, Smith TPL, Medrano JF (2020) De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 9:giaa021

150. National Center for Biotechnology Information (2018) ARS-UCD1.2 Assembly. https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1. Accessed 31 May 2018

151. Broad Institute (2018) Picard Tools version 2.18.2. In: Broad Institute, GitHub Repos. http://broadinstitute.github.io/picard/. Accessed 20 Aug 2018

152. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

153. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498

154. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nat Biotechnol 29:24–26

155. National Center for Biotechnology Information (2018) NCBI Bos taurus Annotation Release 106. https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bos_taurus/106/. Accessed 31 Aug 2018

156. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6:80–92

157. Consortium BGS and A, Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, Weinstock GM, Adelson DL, Eichler EE, Elnitski L, Guigó R, Hamernik DL, Kappes SM, Lewin HA, Lynn DJ, Nicholas FW, Reymond A, Rijnkels M, Skow LC, Zdobnov EM, Schook L, Womack J, Alioto T, Antonarakis SE, Astashyn A, Chapple CE, Chen H-C, Chrast J, Câmara F, Ermolaeva O, Henrichsen CN, Hlavina W, Kapustin Y, Kiryutin B, Kitts P, Kokocinski F, Landrum M, Maglott D, Pruitt K, Sapojnikov V, Searle SM, Solovyev V, Souvorov A, Ucla C, Wyss C, Anzola JM, Gerlach D, Elhaik E, Graur D, Reese JT, Edgar RC, McEwan JC, Payne GM, Raison JM, Junier T, Kriventseva E V, Eyras E, Plass M, Donthu R, Larkin DM, Reecy J, Yang MQ, Chen L, Cheng Z, Chitko-McKown CG, Liu GE, Matukumalli LK, Song J, Zhu B, Bradley DG, Brinkman FSL, Lau LPL, Whiteside MD, Walker A, Wheeler TT, Casey T, German JB, Lemay DG, Maqbool NJ, Molenaar AJ, Seo S, Stothard P, Baldwin CL, Baxter R, Brinkmeyer-Langford CL, Brown WC, Childers CP, Connelley T, Ellis SA, Fritz K, Glass EJ, Herzig CTA, Iivanainen A, Lahmers KK, Bennett AK, Dickens CM, Gilbert JGR, Hagen DE, Salih H, Aerts J, Caetano AR, Dalrymple B, Garcia JF, Gill CA, Hiendleder SG, Memili E, Spurlock D, Williams JL, Alexander L, Brownstein MJ, Guan L, Holt RA, Jones SJM, Marra MA, Moore R, Moore SS, Roberts

A, Taniguchi M, Waterman RC, Chacko J, Chandrabose MM, Cree A, Dao MD, Dinh HH, Gabisi RA, Hines S, Hume J, Jhangiani SN, Joshi V, Kovar CL, Lewis LR, Liu Y-S, Lopez J, Morgan MB, Nguyen NB, Okwuonu GO, Ruiz SJ, Santibanez J, Wright RA, Buhay C, Ding Y, Dugan-Rocha S, Herdandez J, Holder M, Sabo A, Egan A, Goodell J, Wilczek-Boney K, Fowler GR, Hitchens ME, Lozado RJ, Moen C, Steffen D, Warren JT, Zhang J, Chiu R, Schein JE, Durbin KJ, Havlak P, Jiang H, Liu Y, Qin X, Ren Y, Shen Y, Song H, Bell SN, Davis C, Johnson AJ, Lee S, Nazareth L V, Patel BM, Pu L-L, Vattathil S, Williams Jr RL, Curry S, Hamilton C, Sodergren E, Wheeler DA, Barris W, Bennett GL, Eggen A, Green RD, Harhay GP, Hobbs M, Jann O, Keele JW, Kent MP, Lien S, McKay SD, McWilliam S, Ratnakumar A, Schnabel RD, Smith T, Snelling WM, Sonstegard TS, Stone RT, Sugimoto Y, Takasuga A, Taylor JF, Van Tassell CP, Macneil MD, Abatepaulo ARR, Abbey CA, Ahola V, Almeida IG, Amadio AF, Anatriello E, Bahadue SM, Biase FH, Boldt CR, Carroll JA, Carvalho WA, Cervelatti EP, Chacko E, Chapin JE, Cheng Y, Choi J, Colley AJ, de Campos TA, De Donato M, Santos IKF de M, de Oliveira CJF, Deobald H, Devinoy E, Donohue KE, Dovc P, Eberlein A, Fitzsimmons CJ, Franzin AM, Garcia GR, Genini S, Gladney CJ, Grant JR, Greaser ML, Green JA, Hadsell DL, Hakimov HA, Halgren R, Harrow JL, Hart EA, Hastings N, Hernandez M, Hu Z-L, Ingham A, Iso-Touru T, Jamis C, Jensen K, Kapetis D, Kerr T, Khalil SS, Khatib H, Kolbehdari D, Kumar CG, Kumar D, Leach R, Lee JC-M, Li C, Logan KM, Malinverni R, Marques E, Martin WF, Martins NF, Maruyama SR, Mazza R, McLean KL, Medrano JF, Moreno BT, Moré DD, Muntean CT, Nandakumar HP, Nogueira MFG, Olsaker I, Pant SD, Panzitta F, Pastor RCP, Poli MA, Poslusny N, Rachagani S, Ranganathan S, Razpet A, Riggs PK, Rincon G, Rodriguez-Osorio N, Rodriguez-Zas SL, Romero NE, Rosenwald A, Sando L, Schmutz SM, Shen L, Sherman L, Southey BR, Lutzow YS, Sweedler J V, Tammen I, Telugu BPVL, Urbanski JM, Utsunomiya YT, Verschoor CP, Waardenberg AJ, Wang Z, Ward R, Weikard R, Welsh Jr TH, White SN, Wilming LG, Wunderlich KR, Yang J, Zhao F-Q (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528

158. National Center for Biotechnology Information (NCBI) (1988) Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/. Accessed 9 Nov 2021

159. Wikimedia commons (2010) Dominette. https://commons.wikimedia.org/wiki/File:Hereford67-300.jpg. Accessed 9 Dec 2019

160. Koren S, Rhie A, Walenz BP, Dilthey AT, Bickhart DM, Kingan SB, Hiendleder S, Williams JL, Smith TPL, Phillippy AM (2018) De novo assembly of haplotype-resolved genomes with trio binning. Nat Biotechnol 36:1174–1182

161. Heaton MP, Smith TPL, Bickhart DM, Vander Ley BL, Kuehn LA, Oppenheimer J, Shafer WR, Schuetze FT, Stroud B, McClure JC, Barfield JP, Blackburn HD, Kalbfleisch TS, Davenport KM, Kuhn KL, Green RE, Shapiro B, Rosen BD (2021) A Reference Genome Assembly of Simmental

Cattle, Bos taurus taurus. J Hered 112:184–191

162. Rice ES, Koren S, Rhie A, Heaton MP, Kalbfleisch TS, Hardy T, Hackett PH, Bickhart DM, Rosen BD, Ley B Vander, Maurer NW, Green RE, Phillippy AM, Petersen JL, Smith TPL (2020) Continuous chromosome-scale haplotypes assembled from a single interspecies F1 hybrid of yak and cattle. Gigascience 9:giaa029

163. Oppenheimer J, Rosen BD, Heaton MP, Vander Ley BL, Shafer WR, Schuetze FT, Stroud B, Kuehn LA, McClure JC, Barfield JP, Blackburn HD, Kalbfleisch TS, Bickhart DM, Davenport KM, Kuhn KL, Green RE, Shapiro B, Smith TPL (2021) A reference genome assembly of American Bison, Bison bison bison. J Hered 112:174–183

164. Häfliger IM, Marchionatti E, Stengård M, Wolf-Hofstetter S, Paris JM, Jacinto JGP, Watté C, Voelter K, Occelli LM, Komáromy AM, Oevermann A, Goepfert C, Borgo A, Roduit R, Spengeler M, Seefried FR, Drögemüller C (2021) CNGB3 missense variant causes recessive achromatopsia in Original Braunvieh cattle. Int J Mol Sci 22:12440

165. Häfliger IM, Hofstetter S, Mock T, Stettler MH, Meylan M, Mehinagic K, Stokar-Regenscheit N, Drögemüller C (2019) APOB-associated cholesterol deficiency in Holstein cattle is not a simple recessive disease. Anim Genet 50:372–375

166. Jacinto JGP, Häfliger IM, Bernardini M, Mandara MT, Bianchi E, Bolcato M, Romagnoli N, Gentile A, Drögemüller C (2021) A homozygous missense variant in laminin subunit beta 1 as candidate causal mutation of hemifacial microsomia in Romagnola cattle. J Vet Intern Med. https://doi.org/10.1111/jvim.16316

167. O'toole D, Häfliger IM, Leuthard F, Schumaker B, Steadman L, Murphy B, Drögemüller C, Leeb T (2021) X-linked hypohidrotic ectodermal dysplasia in crossbred beef cattle due to a large deletion in eda. Animals 11:657

168. Jacinto JGP, Häfliger IM, Veiga I, Letko A, Gentile A, Drögemüller C (2021) A frameshift insertion in FA2H causes a recessively inherited form of ichthyosis congenita in Chianina cattle. Mol Genet Genomics 296:1313–1322

169. Häfliger IM, Koch CT, Michel A, Rüfenacht S, Meylan M, Welle MM, Drögemüller C (2022) DSP missense variant in a Scottish Highland calf with congenital ichthyosis, alopecia, acantholysis of the tongue and corneal defects. BMC Vet Res 18:20

170. Häfliger IM, Sickinger M, Holsteg M, Raeder LM, Henrich M, Marquardt S, Drögemüller C, Lühken G (2020) An IL17RA frameshift variant in a Holstein cattle family with psoriasis-like skin alterations and immunodeficiency. BMC Genet 21:1–10

171. Häfliger IM, Wiedemar N, Švara T, Starič J, Cociancich V, Šest K, Gombač M, Paller T, Agerholm JS, Drögemüller C (2020) Identification of small and large genomic candidate variants in bovine pulmonary hypoplasia and anasarca syndrome. Anim Genet 51:382–290

172. Häfliger IM, Agerholm JS, Drögemüller C (2020) Constitutional trisomy 20

in an aborted Holstein fetus with pulmonary hypoplasia and anasarca syndrome. Anim. Genet.

173. Letko A, Strugnell B, Häfliger IM, Paris JM, Waine K, Drögemüller C, Scholes S (2021) Compound heterozygous PLA2G6 loss-of-function variants in Swaledale sheep with neuroaxonal dystrophy. Mol Genet Genomics 296:235–242

174. Dürig N, Letko A, Lepori V, Hadji Rasouliha S, Loechel R, Kehl A, Hytönen MK, Lohi H, Mauri N, Dietrich J (2018) Two MC1R loss-of-function alleles in cream-coloured Australian Cattle Dogs and white Huskies. Anim Genet 49:284–290

175. Häfliger IM, Wolf-Hofstetter S, Casola C, Hetzel U, Seefried FR, Drögemüller C (2022) A de novo variant in the bovine ADAMTSL4 gene in an Original Braunvieh calf with congenital cataract. Anim Genet 00:1–6

176. Häfliger IM, Behn H, Freick M, Jagannathan V, Drögemüller C (2019) A COL2A1 de novo variant in a Holstein bulldog calf. Anim Genet 50:113–114

177. Jacinto JGP, Häfliger IM, Gentile A, Drögemüller C, Bolcato M (2021) A 6.7 kb deletion in the COL2A1 gene in a Holstein calf with achondrogenesis type II and perosomus elumbis. Anim Genet 52:244–245

178. Jacinto JGP, Häfliger IM, Letko A, Drögemüller C, Agerholm JS (2020) A large deletion in the COL2A1 gene expands the spectrum of pathogenic variants causing bulldog calf syndrome in cattle. Acta Vet Scand 62:1–8

179. Häfliger IM, Letko A, Murgiano L, Drögemüller C (2020) De novo stop-lost germline mutation in FGFR3 causes severe chondrodysplasia in the progeny of a Holstein bull. Anim Genet 51:466–469

180. Jacinto JGP, Häfliger IM, Veiga IMB, Letko A, Benazzi C, Bolcato M, Drögemüller C (2020) A heterozygous missense variant in the COL5A2 in holstein cattle resembling the classical Ehlers–Danlos syndrome. Animals 10:2002

181. Jacinto JGP, Häfliger IM, Akyürek EE, Sacchetto R, Benazzi C, Gentile A, Drögemüller C (2021) KCNG1-related syndromic form of congenital neuromuscular channelopathy in a crossbred calf. Genes (Basel) 12:1792

182. Jacinto JGP, Häfliger IM, Veiga IMB, Drögemüller C, Agerholm JS (2020) A de novo mutation in KRT5 in a crossbred calf with epidermolysis bullosa simplex. J Vet Intern Med 34:2800–2807

183. Jacinto JGP, Häfliger IM, McEvoy FJ, Drögemüller C, Agerholm JS (2021) A de novo mutation in COL1A1 in a Holstein Calf with Osteogenesis Imperfecta Type II. Animals 11:561

184. Häfliger IM, Seefried F, Drögemüller C (2020) Trisomy 29 in a stillborn Swiss Original Braunvieh calf. Anim. Genet.

185. Jacinto JGP, Häfliger IM, Gentile A, Drögemüller C (2021) A heterozygous missense variant in MAP2K2 in a stillborn romagnola calf with skeletal-cardio-enteric dysplasia. Animals 11:1931

186. Laffleur B, Lim J, Zhang W, Chen Y, Pefanis E, Bizarro J, Batista CR, Wu

L, Economides AN, Wang J, Basu U (2021) Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. Nat Genet 53:230–242

187. Cheong A, Lingutla R, Mager J (2020) Expression analysis of mammalian mitochondrial ribosomal protein genes. Gene Expr Patterns 38:119147

188. Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, Claas FH, Walker JJ, Redman CW, Morgan L, Tower C, Regan L, Moore GE, Carrington M, Moffett A (2010) Maternal activating KIRs protect against human reproductive failure mediated by fetal HLA-C2. J Clin Invest 120:4102–4110

189. Häfliger IM, Seefried FR, Spengeler M, Drögemüller C (2021) Mining massive genomic data of two Swiss Braunvieh cattle populations reveals six novel candidate variants that impair reproductive success. Genet Sel Evol 53:1–16

190. Häfliger IM, Spengeler M, Seefried FR, Drögemüller C (2022) Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle. Sci Rep 12:1–13

191. Häfliger IM, Seefried FR, Drögemüller C (2021) Reverse genetic screen for deleterious recessive variants in the local Simmental cattle population of Switzerland. Animals 11:3535

192. Reynolds EGM, Neeley C, Lopdell TJ, Keehan M, Dittmer K, Harland CS, Couldrey C, Johnson TJJ, Tiplady K, Worth G, Walker M, Davis SR, Sherlock RG, Carnie K, Harris BL, Charlier C, Georges M, Spelman RJ, Garrick DJ, Littlejohn MD (2021) Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. Nat Genet 53:949–954

193. Kadri NK, Sahana G, Charlier C, Iso-Touru T, Guldbrandtsen B, Karim L, Nielsen US, Panitz F, Aamand GP, Schulman N, Georges M, Vilkki J, Lund MS, Druet T (2014) A 660-kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic Red cattle: additional evidence for the common occurrence of balancing selection in livestock. PLoS Genet 10:e1004049

194. Pacheco HA, Rezende FM, Peñagaricano F (2020) Gene mapping and genomic prediction of bull fertility using sex chromosome markers. J Dairy Sci 103:3304–3311

195. Zhou J, Troyanskaya OG (2015) Predicting effects of noncoding variants with deep learning-based sequence model. Nat Methods 12:931–934

196. Kadri NK, Mapel XM, Pausch H (2021) The intronic branch point sequence is under strong evolutionary constraint in the bovine and human genome. Commun Biol 4:1206

197. Giuffra E, Tuggle CK, Archibald AL, Acloque H, Chamberlain AJ, Cochrane G, Daetwyler HD, Djebali S, Eory L, Foissac S, Goddard ME, Groenen MAM, Halstead M, Harrison PW, Hayes BJ, Kern C, Khatib H, Kuo RI, MacHugh DE, Macleod IM, Madsen O, Reecy JM, Ross PJ, Watson M, Wells JM, Zhou HJ, Consortium F (2019) Functional Annotation of Animal Genomes (FAANG): Current achievements and roadmap. Annu Rev Anim

Biosci 7:65–88

198. Lavrichenko K, Johansson S, Jonassen I (2021) Comprehensive characterization of copy number variation (CNV) called from array, long- and short-read data. BMC Genomics 22:826

199. Iannuzzi A, Parma P, Iannuzzi L (2021) Chromosome abnormalities and fertility in domestic bovids: A review. Animals 11:802

200. Lloret-Villas A, Bhati M, Kadri NK, Fries R, Pausch H (2021) Investigating the impact of reference assembly choice on genomic analyses in a cattle breed. BMC Genomics 22:363

201. Crysnanto D, Pausch H (2020) Bovine breed-specific augmented reference graphs facilitate accurate sequence read mapping and unbiased variant discovery. Genome Biol 21:184

202. Crysnanto D, Leonard AS, Fang Z-H, Pausch H (2021) Novel functional sequences uncovered through a bovine multiassembly graph. Proc Natl Acad Sci 118:e2101056118

203. Leonard AS, Crysnanto D, Fang Z-H, Heaton MP, Vander Ley BL, Herrera C, Bollwein H, Bickhart DM, Kuhn KL, Smith TPL, Rosen BD, Pausch H (2021) Bovine pangenome reveals trait-associated structural variation from diverse assembly inputs. bioRxiv 2021.11.02.466900

204. Jacinto JGP, Häfliger IM, Borel N, Zanolari P, Drögemüller C, Veiga I (2021) Clinicopathological and genomic characterization of a Simmental calf with generalized bovine juvenile angiomatosis. Animals 11:624

205. Brunner MAT, Veiga IMB, Niggeler A, Häfliger IM, Stettler M, Meylan M, Welle M, Drögemüller C (2020) Is a de novo nonsense variant in the ASPDH gene the cause of ulcerative skin lesions in a Holstein calf? Vet Dermatol 31:244-e54

206. Havrilla JM, Pedersen BS, Layer RM, Quinlan AR (2019) A map of constrained coding regions in the human genome. Nat Genet 51:88–95

207. 100 000 Genomes Project Pilot Investigators (2021) 100,000 Genomes pilot on rare-disease diagnosis in health care — preliminary report. N Engl J Med 385:1868–1880

208. Derks MFL, Steensma M (2021) Review: Balancing Selection for Deleterious Alleles in Livestock. Front Genet 12:2526

209. Hjortø L, Henryon M, Liu H, Berg P, Thomasen JR, Sørensen AC (2021) Pre-selection against a lethal recessive allele in breeding schemes with optimum-contribution selection or truncation selection. Genet Sel Evol 53:75

210. Johnsson M, Gaynor RC, Jenko J, Gorjanc G, de Koning D-J, Hickey JM (2019) Removal of alleles by genome editing (RAGE) against deleterious load. Genet Sel Evol 51:14

211. Gim G, Kwon D, Eom K, Moon J, Park J, Lee W, Jung D, Kim D, Yi J, Ha J (2021) Production of MSTN-mutated cattle without exogenous gene integration using CRISPR-Cas9. Biotechnol J 2100198

212. Laible G, Cole S-A, Brophy B, Wei J, Leath S, Jivanji S, Littlejohn MD, Wells

DN (2021) Holstein Friesian dairy cattle edited for diluted coat color as a potential adaptation to climate change. BMC Genomics 22:856

213. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N (2011) Highly efficient modification of beta-lactoglobulin (BLG) gene via zinc-finger nucleases in cattle. Cell Res 21:1638–1640

214. Carlson DF, Lancto CA, Zang B, Kim E-S, Walton M, Oldeschulte D, Seabury C, Sonstegard TS, Fahrenkrug SC (2016) Production of hornless dairy cattle from genome-edited cell lines. Nat Biotechnol 34:479–481

215. Van Eenennaam AL (2019) Application of genome editing in farm animals: cattle. Transgenic Res 28:93–100

216. Eager KLM, Conyers LE, Woolley SA, Tammen I, O'Rourke BA (2020) A novel ABCA12 frameshift mutation segregates with ichthyosis fetalis in a Polled Hereford calf. Anim Genet 51:837–838

217. Woolley SA, Eager KLM, Häfliger IM, Bauer A, Drögemüller C, Leeb T, O'Rourke BA, Tammen I (2019) An ABCA12 missense variant in a Shorthorn calf with ichthyosis fetalis. Anim Genet 50:749–752

218. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, Carta E, Dardano S, Dive M, Fasquelle C, Frennet J-C, Hanset R, Hubin X, Jorgensen C, Karim L, Kent M, Harvey K, Pearce BR, Simon P, Tama N, Nie H, Vandeputte S, Lien S, Longeri M, Fredholm M, Harvey RJ, Georges M (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. Nat Genet 40:449–454

219. Pausch H, Ammermuller S, Wurmser C, Hamann H, Tetens J, Drogemuller C, Fries R (2016) A nonsense mutation in the COL7A1 gene causes epidermolysis bullosa in Vorderwald cattle. BMC Genet 17:149

220. Menoud A, Welle M, Tetens J, Lichtner P, Drögemüller C (2012) A COL7A1 mutation causes dystrophic epidermolysis bullosa in Rotes Höhenvieh cattle. PLoS One 7:e38823

221. Hofstetter S, Welle M, Gorgas D, Balmer P, Roosje P, Mock T, Meylan M, Jagannathan V, Drögemüller C (2017) A de novo germline mutation of DLX 3 in a Brown Swiss calf with tricho-dento-osseus-like syndrome. Vet Dermatol 28:616-e150

222. Drögemüller C, Distl O, Leeb T (2003) X-linked anhidrotic ectodermal dysplasia (ED1) in men, mice, and cattle. Genet Sel Evol 35:S137

223. Gargani M, Valentini A, Pariset L (2011) A novel point mutation within the EDA gene causes an exon dropping in mature RNA in Holstein Friesian cattle breed affected by X-linked anhidrotic ectodermal dysplasia. BMC Vet Res 7:35

224. Ogino A, Kohama N, Ishikawa S, Tomita K, Nonaka S, Shimizu K, Tanabe Y, Okawa H, Morita M (2011) A novel mutation of the bovine EDA gene associated with anhidrotic ectodermal dysplasia in Holstein cattle. Hereditas 148:46–49

225. Ogino A, Shimizu K, Tanabe Y, Morita M (2012) De novo mutation of the bovine EDA gene associated with anhidrotic ectodermal dysplasia in Japanese Black cattle. Anim Genet 43:646

226. Escouflaire C, Rebours E, Charles M, Orellana S, Cano M, Rivière J, Grohs C, Hayes H, Capitan A (2019) A de novo 3.8-Mb inversion affecting the EDA and XIST genes in a heterozygous female calf with generalized hypohidrotic ectodermal dysplasia. BMC Genomics 20:715

227. Karlskov-Mortensen P, Cirera S, Nielsen OL, Arnbjerg J, Reibel J, Fredholm M, Agerholm JS (2011) Exonization of a LINE1 fragment implicated in X-linked hypohidrotic ectodermal dysplasia in cattle. Anim Genet 42:578–584

228. Drögemüller C, Barlund CS, Palmer CW, Leeb T (2006) A novel mutation in the bovine EDA gene causing anhidrotic ectodermal dysplasia (Brief report). Arch Anim Breed 49:615–616

229. Drögemüller C, Distl O, Leeb T (2001) Partial deletion of the bovine ED1 gene causes anhidrotic ectodermal dysplasia in cattle. Genome Res 11:1699–1705

230. Drögemüller C, Peters M, Pohlenz J, Distl O, Leeb T (2002) A single point mutation within the ED1 gene disrupts correct splicing at two different splice sites and leads to anhidrotic ectodermal dysplasia in cattle. J Mol Med 80:319–323

231. Tajima M, Miyake S, Takehana K, Kobayashi A, Yamato O, Maede Y (1999) Gene defect of dermatan sulfate proteoglycan of cattle affected with a variant form of Ehlers-Danlos syndrome. J Vet Intern Med 13:202–205

232. Kuca T, Marron BM, Jacinto JGP, Paris JM, Gerspach C, Beever JE, Drögemüller C (2021) A nonsense variant in hephaestin like 1 (HEPHL1) is responsible for congenital hypotrichosis in Belted Galloway cattle. Genes (Basel) 12:643

233. Michot P, Fantini O, Braque R, Allais-Bonnet A, Saintilan R, Grohs C, Barbieri J, Genestout L, Danchin-Burge C, Gourreau J-M, Boichard D, Pin D, Capitan A (2015) Whole-genome sequencing identifies a homozygous deletion encompassing exons 17 to 23 of the integrin beta 4 gene in a Charolais calf with junctional epidermolysis bullosa. Genet Sel Evol 47:37

234. Peters M, Reber I, Jagannathan V, Raddatz B, Wohlsein P, Drögemüller C (2015) DNA-based diagnosis of rare diseases in veterinary medicine: a 4.4 kb deletion of ITGB4 is associated with epidermolysis bullosa in Charolais cattle. BMC Vet Res 11:1–8

235. Ford CA, Stanfield AM, Spelman RJ, Smits B, Ankersmidt-Udy AEL, Cottier K, Holloway H, Walden A, Al-Wahb M, Bohm E, Snell RG, Sutherland GT (2005) A mutation in bovine keratin 5 causing epidermolysis bullosa simplex, transmitted by a mosaic sire. J Invest Dermatol 124:1170–1176

236. Jacinto JGP, Markey AD, Veiga IMB, Paris JM, Welle M, Beever JE, Drögemüller C (2021) A KRT71 loss-of-function variant results in inner root sheath dysplasia and recessive congenital hypotrichosis of Hereford cattle. Genes (Basel) 12:1038

237. Sartelet A, Harland C, Tamma N, Karim L, Bayrou C, Li W, Ahariz N, Coppieters W, Georges M, Charlier C (2015) A stop-gain in the laminin, alpha 3 gene causes recessive junctional epidermolysis bullosa in Belgian

Blue cattle. Anim Genet 46:566–570

238. Murgiano L, Wiedemar N, Jagannathan V, Isling LK, Drögemüller C, Agerholm JS (2015) Epidermolysis bullosa in Danish Hereford calves is caused by a deletion in LAMC2 gene. BMC Vet Res 11:23

239. Knaust J, Hadlich F, Weikard R, Kuehn C (2016) Epistatic interactions between at least three loci determine the “rat-tail” phenotype in cattle. Genet Sel Evol 48:26

240. Jung S, Pausch H, Langenmayer MC, Schwarzenbacher H, Majzoub-Altweck M, Gollnick NS, Fries R (2014) A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. BMC Genomics 15:623

241. Murgiano L, Shirokova V, Welle MM, Jagannathan V, Plattet P, Oevermann A, Pienkowska-Schelling A, Gallo D, Gentile A, Mikkola M, Drögemüller C (2015) Hairless streaks in cattle implicate TSR2 in early hair follicle formation. PLOS Genet 11:e1005427

242. Struck A-K, Dierks C, Braun M, Hellige M, Wagner A, Oelmaier B, Beineke A, Metzger J, Distl O (2018) A recessive lethal chondrodysplasia in a miniature zebu family results from an insertion affecting the chondroitin sulfat domain of aggrecan. BMC Genet 19:1–9

243. Brenig B, Schütz E, Hardt M, Scheuermann P, Freick M (2015) A 20 bp duplication in exon 2 of the aristaless-like homeobox 4 gene (ALX4) Is the candidate causative mutation for tibial hemimelia syndrome in Galloway cattle. PLoS One 10:e0129208

244. Beever JE, Marron BM (2012) Screening for the genetic defect causing tibial hemimelia in bovines. U.S. Pat. No. 8,158,356

245. Floriot S, Vesque C, Rodriguez S, Bourgain-Guglielmetti F, Karaiskou A, Gautier M, Duchesne A, Barbey S, Fritz S, Vasilescu A, Bertaud M, Moudjou M, Halliez S, Cormier-Daire V, E.L. Hokayem J, Nigg EA, Manciaux L, Guatteo R, Cesbron N, Toutirais G, Eggen A, Schneider-Maunoury S, Boichard D, Sobczak-Thépot J, Schibler L (2015) C-Nap1 mutation affects centriole cohesion and is associated with a Seckel-like syndrome in cattle. Nat Commun 6:6894

246. Petersen JL, Tietze SM, Burrack RM, Steffen DJ (2019) Evidence for a de novo, dominant germ-line mutation causative of osteogenesis imperfecta in two Red Angus calves. Mamm Genome 30:81–87

247. Reinartz S, Mohwinkel H, Surie C, Hellige M, Feige K, Eikelberg D, Beineke A, Metzger J, Distl O (2017) Germline mutation within COL2A1 associated with lethal chondrodysplasia in a polled Holstein family. BMC Genomics 18:1–10

248. Agerholm JS, Menzi F, McEvoy FJ, Jagannathan V, Drogemuller C (2016) Lethal chondrodysplasia in a family of Holstein cattle is associated with a de novo splice site variant of COL2A1. BMC Vet Res 12:1–9

249. Sieck RL, Fuller AM, Bedwell PS, Ward JA, Sanders SK, Xiang S-H, Peng S, Petersen JL, Steffen DJ (2020) Mandibulofacial dysostosis attributed to a recessive mutation of CYP26C1 in Hereford cattle. Genes (Basel)

11:1246

250. Agerholm JS, McEvoy F, Arnbjerg J (2006) Brachyspina syndrome in a Holstein calf. J Vet Diagnostic Investig 18:418–422

251. Charlier C, Agerholm JS, Coppieters W, Karlskov-Mortensen P, Li W, de Jong G, Fasquelle C, Karim L, Cirera S, Cambisano N, Ahariz N, Mullaart E, Georges M, Fredholm M (2012) A deletion in the bovine FANCI gene compromises fertility by causing fetal death and brachyspina. PLoS One 7:e43085

252. Singleton AC, Mitchell AL, Byers PH, Potter KA, Pace JM (2005) Bovine model of Marfan syndrome results from an amino acid change (c. 3598G> A, p. E1200K) in a calcium-binding epidermal growth factor-like domain of fibrillin-1. Hum Mutat 25:348–352

253. Hirano T, Matsuhashi T, Kobayashi N, Watanabe T, Sugimoto Y (2012) Identification of an FBN1 mutation in bovine Marfan syndrome-like disease. Anim Genet 43:11–17

254. Agerholm JS, McEvoy FJ, Heegaard S, Charlier C, Jagannathan V, Drögemüller C (2017) A de novo missense mutation of FGFR2 causes facial dysplasia syndrome in Holstein cattle. BMC Genet 18:1–9

255. Schwarzenbacher H, Wurmser C, Flisikowski K, Misurova L, Jung S, Langenmayer MC, Schnieke A, Knubben-Schweizer G, Fries R, Pausch H (2016) A frameshift mutation in GON4L is associated with proportionate dwarfism in Fleckvieh cattle. Genet Sel Evol 48:1–10

256. Seichter D, Russ I, Förster M, Medugorac I (2011) SNP-based association mapping of Arachnomelia in Fleckvieh cattle. Anim Genet 42:544–547

257. Buitkamp J, Semmer J, Götz K-U (2011) Arachnomelia syndrome in Simmental cattle is caused by a homozygous 2-bp deletion in the molybdenum cofactor synthesis step 1 gene (MOCS1). BMC Genet 12:11

258. Fasquelle C, Sartelet A, Li W, Dive M, Tamma N, Michaux C, Druet T, Huijbers IJ, Isacke CM, Coppieters W, Georges M, Charlier C (2009) Balancing selection of a frame-shift mutation in the MRC2 gene accounts for the outbreak of the crooked tail syndrome in Belgian Blue cattle. PLOS Genet 5:e1000666

259. Sartelet A, Klingbeil P, Franklin CK, Fasquelle C, Geron S, Isacke CM, Georges M, Charlier C (2012) Allelic heterogeneity of Crooked Tail Syndrome: result of balancing selection? Anim Genet 43:604–607

260. Koltes JE, Mishra BP, Kumar D, Kataria RS, Totir LR, Fernando RL, Cobbold R, Steffen D, Coppieters W, Georges M (2009) A nonsense mutation in cGMP-dependent type II protein kinase (PRKG2) causes dwarfism in American Angus cattle. Proc Natl Acad Sci 106:19250–19255

261. Sartelet A, Druet T, Michaux C, Fasquelle C, Géron S, Tamma N, Zhang Z, Coppieters W, Georges M, Charlier C (2012) A splice site variant in the bovine RNF11 gene compromises growth and regulation of the inflammatory response. PLoS Genet 8:e1002581–e1002581

262. Meyers SN, McDaneld TG, Swist SL, Marron BM, Steffen DJ, O'Toole D, O'Connell JR, Beever JE, Sonstegard TS, Smith TPL (2010) A deletion

mutation in bovine SLC4A2 is associated with osteopetrosis in Red Angus cattle. BMC Genomics 11:337

263. Thomsen B, Horn P, Panitz F, Bendixen E, Petersen AH, Holm L-E, Nielsen VH, Agerholm JS, Arnbjerg J, Bendixen C (2006) A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. Genome Res 16:97–105

264. Agerholm JS, Bendixen C, Andersen O, Arnbjerg J (2001) Complex vertebral malformation in Holstein calves. J Vet Diagnostic Investig 13:283–289

265. Yuzbasiyan-Gurkan V, Bartlett E (2006) Identification of a unique splice site variant in SLC39A4 in bovine hereditary zinc deficiency, lethal trait A46: An animal model of acrodermatitis enteropathica. Genomics 88:521–526

266. Drögemüller C, Tetens J, Sigurdsson S, Gentile A, Testoni S, Lindblad-Toh K, Leeb T (2010) Identification of the bovine arachnomelia mutation by massively parallel sequencing implicates sulfite oxidase (SUOX) in bone development. PLOS Genet 6:e1001079

267. Kromik A, Ulrich R, Kusenda M, Tipold A, Stein VM, Hellige M, Dziallas P, Hadlich F, Widmann P, Goldammer T, Baumgärtner W, Rehage J, Segelke D, Weikard R, Kühn C (2015) The mammalian cervical vertebrae blueprint depends on the T (brachyury) gene. Genetics 199:873–883

268. Braun M, Lehmbecker A, Eikelberg D, Hellige M, Beineke A, Metzger J, Distl O (2021) De novo ZIC2 frameshift variant associated with frontonasal dysplasia in a Limousin calf. BMC Genomics 22:1–9

269. Denholm LJ, Beever JE, Marron BM, Walker KH, Healy PJ, Dennis JA, O'Rourke BA, Teseling TF (2014) Contractural Arachnodactyly (CA): A congenital recessive marfanoid syndrome with reduced elasticity of muscle connective tissue resulting from a large (~58 kilobase pair) deletion affecting the ADAMTSL3 gene in Angus and Angus influenced cattle breeds. In: Proc. XXVIII World Buiatrics Congr. Cairns Aust. p 0305

270. Beever JE, Marron BM (2011) Screening for arthrogryposis multiplex in bovines. U.S. Pat. No. 20,110,151,440 A1

271. Dennis JA, Healy PJ, Beaudet AL, O'Brien WE (1989) Molecular definition of bovine argininosuccinate synthetase deficiency. Proc Natl Acad Sci USA 86:7947–7951

272. Drögemüller C, Drögemüller M, Leeb T, Mascarello F, Testoni S, Rossi M, Gentile A, Damiani E, Sacchetto R (2008) Identification of a missense mutation in the bovine ATP2A1 gene in congenital pseudomyotonia of Chianina cattle: An animal model of human Brody disease. Genomics 92:474–477

273. Murgiano L, Sacchetto R, Testoni S, Dorotea T, Mascarello F, Liguori R, Gentile A, Drögemüller C (2012) Pseudomyotonia in Romagnola cattle caused by novel ATP2A1 mutations. BMC Vet Res 8:1–9

274. DENNIS JA, HEALY PJ (1999) Definition of the mutation responsible for maple syrup urine disease in poll shorthorns and genotyping poll

shorthorns and poll herefords for maple syrup urine disease alleles. Res Vet Sci 67:1–6

275. Zhang B, Healy PJ, Zhao Y, Crabb DW, Harris RA (1990) Premature translation termination of the pre-E1 alpha subunit of the branched chain alpha-ketoacid dehydrogenase as a cause of maple syrup urine disease in Polled Hereford calves. J Biol Chem 265:2425–2427

276. Agerholm JS, McEvoy FJ, Menzi F, Jagannathan V, Drögemüller C (2016) A CHRNB1 frameshift mutation is associated with familial arthrogryposis multiplex congenita in Red dairy cattle. BMC Genomics 17:1–10

277. Kraner S, Sieb JP, Thompson PN, Steinlein OK (2002) Congenital myasthenia in Brahman calves caused by homozygosity for a CHRNE truncating mutation. Neurogenetics 4:87–91

278. Houweling PJ, Cavanagh JAL, Palmer DN, Frugier T, Mitchell NL, Windsor PA, Raadsma HW, Tammen I (2006) Neuronal ceroid lipofuscinosis in Devon cattle is caused by a single base duplication (c.662dupG) in the bovine CLN5 gene. Biochim Biophys Acta - Mol Basis Dis 1762:890–897

279. Dennis JA, Moran C, Healy PJ (2000) The bovine α-glucosidase gene: coding region, genomic structure, and mutations that cause bovine generalized glycogenosis. Mamm Genome 11:206–212

280. Pierce KD, Handford CA, Morris R, Vafa B, Dennis JA, Healy PJ, Schofield PR (2001) A nonsense mutation in the α1 subunit of the inhibitory glycine receptor associated with bovine myoclonus. Mol Cell Neurosci 17:354–363

281. El-Hamidi M, Leipold HW, Vestweber JGE, Saperstein G (1989) Spinal muscular atrophy in Brown Swiss calves. Zentralblatt für Vet R A 36:731–738

282. Krebs S, Medugorac I, Röther S, Strässer K, Förster M (2007) A missense mutation in the 3-ketodihydrosphingosine reductase FVT1 as candidate causal mutation for bovine spinal muscular atrophy. Proc Natl Acad Sci 104:6746–6751

283. Duchesne A, Vaiman A, Frah M, Floriot S, Legoueix-Rodriguez S, Desmazières A, Fritz S, Beauvallet C, Albaric O, Venot E, Bertaud M, Saintilan R, Guatteo R, Esquerré D, Branchu J, Fleming A, Brice A, Darios F, Vilotte J-L, Stevanin G, Boichard D, El Hachimi KH (2018) Progressive ataxia of Charolais cattle highlights a role of KIF1C in sustainable myelination. PLOS Genet 14:e1007550

284. Leipprandt JR, Chen H, Horvath JE, Qiao XT, Jones MZ, Friderici KH (1999) Identification of a bovine β-mannosidosis mutation and detection of two β-mannosidase pseudogenes. Mamm Genome 10:1137–1141

285. Tollersrud OK, Berg T, Healy P, Evjen G, Ramachandran U, Nilssen Ø (1997) Purification of bovine lysosomal α-mannosidase, characterization of its gene and determination of two mutations that cause α-mannosidosis. Eur J Biochem 246:410–419

286. Drögemüller C, Reichart U, Seuberlich T, Oevermann A, Baumgartner M, Kühni Boghenbor K, Stoffel MH, Syring C, Meylan M, Müller S, Müller M, Gredler B, Sölkner J, Leeb T (2011) An unusual splice defect in the

mitofusin 2 gene (MFN2) is associated with degenerative axonopathy in Tyrolean Grey cattle. PLoS One 6:e18931

287. Wiedemar N, Riedi A, Jagannathan V, Drögemüller C, Meylan M (2015) Genetic abnormalities in a calf with congenital increased muscular tonus. J Vet Intern Med 29:1418

288. Karageorgos L, Hill B, Bawden MJ, Hopwood JJ (2007) Bovine mucopolysaccharidosis type IIIB. J Inherit Metab Dis 30:358–364

289. Woolley SA, Tsimnadis ER, Lenghaus C, Healy PJ, Walker K, Morton A, Khatkar MS, Elliott A, Kaya E, Hoerner C, Priestman DA, Shepherd D, Platt FM, Porebski BT, Willet CE, O'Rourke BA, Tammen I (2020) Molecular basis for a new bovine model of Niemann-Pick type C disease. PLoS One 15:e0238697

290. Sartelet A, Li W, Pailhoux E, Richard C, Tamma N, Karim L, Fasquelle C, Druet T, Coppieters W, Georges M, Charlier C (2015) Genome-wide next-generation DNA and RNA sequencing reveals a mutation that perturbs splicing of the phosphatidylinositol glycan anchor biosynthesis class H gene (PIGH) and causes arthrogryposis in Belgian Blue cattle. BMC Genomics 16:1–9

291. Leipold HW, Blaugh B, Huston K, Edgerly CG, Hibbs CM (1973) Weaver syndrome in Brown Swiss cattle: clinical signs and pathology. Vet Med small Anim Clin 68:645–647

292. McClure M, Kim E, Bickhart D, Null D, Cooper T, Cole J, Wiggans G, Ajmone-Marsan P, Colli L, Santus E, Liu GE, Schroeder S, Matukumalli L, Van Tassell C, Sonstegard T (2013) Fine mapping for weaver syndrome in Brown Swiss cattle and the identification of 41 concordant mutations across NRCAM, PNPLA8 and CTTNBP2. PLoS One 8:e59251

293. Kunz E, Rothammer S, Pausch H, Schwarzenbacher H, Seefried FR, Matiasek K, Seichter D, Russ I, Fries R, Medugorac I (2016) Confirmation of a non-synonymous SNP in PNPLA8 as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. Genet Sel Evol 48:1–14

294. Richt JA, Hall SM (2008) BSE Case Associated with Prion Protein Gene Mutation. PLOS Pathog 4:e1000156

295. Tsujino S, Shanske S, Valberg SJ, Cardinet GH, Smith BP, DiMauro S (1996) Cloning of bovine muscle glycogen phosphorylase cDNA and identification of a mutation in cattle with myophosphorylase deficiency, an animal model for McArdle's disease. Neuromuscul Disord 6:19–26

296. Becker D, Weikard R, Schulze C, Wohlsein P, Kühn C (2020) A 50-kb deletion disrupting the RSPO2 gene is associated with tetradysmelia in Holstein Friesian cattle. Genet Sel Evol 52:1–13

297. Duchesne A, Vaiman A, Castille J, Beauvallet C, Gaignard P, Floriot S, Rodriguez S, Vilotte M, Boulanger L, Passet B, Albaric O, Guillaume F, Boukadiri A, Richard L, Bertaud M, Timsit E, Guatteo R, Jaffrézic F, Calvel P, Helary L, Mahla R, Esquerré D, Péchoux C, Liuu S, Vallat J-M, Boichard D, Slama A, Vilotte J-L (2017) Bovine and murine models highlight novel roles for SLC25A46 in mitochondrial dynamics and metabolism, with

implications for human and animal health. PLoS Genet 13:e1006597–e1006597

298. Hafner A, Dahme E, Obermaier G, Schmidt P, Dirksen G (1993) Spinal dysmyelination in new-born Brown Swiss x Braunvieh calves. J Vet Med Ser B 40:413–422

299. Thomsen B, Nissen PH, Agerholm JS, Bendixen C (2010) Congenital bovine spinal dysmyelination is caused by a missense mutation in the SPAST gene. Neurogenetics 11:175–183

300. American Jersey Cattle Association (2020) Jersey neuropathy with splayed forelimbs (JNS). http://www.usjersey.com/portals/0/AJCA/2_Docs/Genetic-Resources/Mailer_FAQ.pdf?utm_source=BenchmarkEmail&utm_campaign=New_Genetic_Abnormality_Factor_Identified_in_Jersey_Breed&utm_medium=email. Accessed 17 Nov 2021

301. Shan S, Xu F, Bleyer M, Becker S, Melbaum T, Wemheuer W, Hirschfeld M, Wacker C, Zhao S, Schütz E, Brenig B (2020) Association of α/β-hydrolase D16B with bovine conception rate and sperm plasma membrane lipid composition. Int J Mol Sci 21:627

302. Iso-Touru T, Wurmser C, Venhoranta H, Hiltpold M, Savolainen T, Sironen A, Fischer K, Flisikowski K, Fries R, Vicente-Carrillo A, Alvarez-Rodriguez M, Nagy S, Mutikainen M, Peippo J, Taponen J, Sahana G, Guldbrandtsen B, Simonen H, Rodriguez-Martinez H, Andersson M, Pausch H (2019) A splice donor variant in CCDC189 is associated with asthenospermia in Nordic Red dairy cattle. BMC Genomics 20:1–13

303. Hirano T, Kobayashi N, Itoh T, Takasuga A, Nakamaru T, Hirotsune S, Sugimoto Y (2000) Null mutation of PCLN-1/Claudin-16 results in bovine chronic interstitial nephritis. Genome Res 10:659–663

304. Hirano T, Hirotsune S, Sasaki S, Kikuchi T, Sugimoto Y (2002) A new deletion mutation in bovine Claudin-16 (CL-16) deficiency and diagnosis. Anim Genet 33:118–122

305. Jenkins MM, LeBoeuf RD, Ruth GR, Bloomer JR (1998) A novel stop codon mutation (X417L) of the ferrochelatase gene in bovine protoporphyria, a natural animal model of the human disease. Biochim Biophys Acta - Mol Basis Dis 1408:18–24

306. Khalaj M, Abbasi AR, Shimojo K, Moritomo Y, Yoneda K, Kunieda T (2009) A missense mutation (p. Leu2153His) of the factor VIII gene causes cattle haemophilia A. Anim Genet 40:763–765

307. Reinartz S, Weiß C, Mischke R, Distl O (2018) A mild form of haemophilia A is associated with two factor VIII missense mutations in German Fleckvieh cattle. Anim Genet 49:350–351

308. Marron BM, Robinson JL, Gentry PA, Beever JE (2004) Identification of a mutation associated with factor XI deficiency in Holstein cattle. Anim Genet 35:454–456

309. Kunieda M, Tsuji T, Abbasi AR, Khalaj M, Ikeda M, Miyadera K, Ogawa H, Kunieda T (2005) An insertion mutation of the bovine F11 gene is

responsible for factor XI deficiency in Japanese black cattle. Mamm Genome 16:383–389

310. Koch CT, Bruggmann R, Tetens J, Drögemüller C (2013) A non-coding genomic duplication at the HMX1 locus is associated with crop ears in Highland cattle. PLoS One 8:e77841

311. Hirano T, Kobayashi N, Matsuhashi T, Watanabe D, Watanabe T, Takasuga A, Sugimoto M, Sugimoto Y (2013) Mapping and exome sequencing identifies a mutation in the IARS gene as the cause of hereditary perinatal weak calf syndrome. PLoS One 8:e64036

312. Shuster DE, Kehrli Jr. ME, Ackermann MR, Gilbert RO (1992) Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. Proc Natl Acad Sci U S A 89:9225–9229

313. Testoni S, Bartolone E, Rossi M, Patrignani A, Bruggmann R, Lichtner P, Tetens J, Gentile A, Drögemüller C (2012) KDM2B is implicated in bovine lethal multi-organic developmental dysplasia. PLoS One 7:e45634

314. Venhoranta H, Pausch H, Wysocki M, Szczerbal I, Hänninen R, Taponen J, Uimari P, Flisikowski K, Lohi H, Fries R, Switonski M, Andersson M (2013) Ectopic KIT copy number variation underlies impaired migration of primordial germ cells associated with gonadal hypoplasia in cattle (Bos taurus). PLoS One 8:e75659–e75659

315. Duchesne A, Gautier M, Chadi S, Grohs C, Floriot S, Gallard Y, Caste G, Ducos A, Eggen A (2006) Identification of a doublet missense substitution in the bovine LRP4 gene as a candidate causal mutation for syndactyly in Holstein cattle. Genomics 88:610–621

316. Eager KLM, Cauchi M, Willet CE, Häfliger IM, Drögemüller C, O'Rourke BA, Tammen I (2021) The previously reported LRP4 c. 4940C> T variant is not associated with syndactyly in cattle. Anim Genet 52:380–381

317. Johnson EB, Steffen DJ, Lynch KW, Herz J (2006) Defective splicing of Megf7/Lrp4, a regulator of distal limb development, in autosomal recessive mulefoot disease. Genomics 88:600–609

318. Drögemüller C, Leeb T, Harlizius B, Tammen I, Distl O, Höltershinken M, Gentile A, Duchesne A, Eggen A (2007) Congenital syndactyly in cattle: four novel mutations in the low density lipoprotein receptor-related protein 4 gene (LRP4). BMC Genet 8:1–12

319. Kunieda T, Nakagiri M, Takami M, Ide H, Ogawa H (1999) Cloning of bovine LYST gene and identification of a missense mutation associated with Chediak-Higashi syndrome of cattle. Mamm Genome 10:1146–1149

320. Flisikowski K, Venhoranta H, Nowacka-Woszuk J, McKay SD, Flyckt A, Taponen J, Schnabel R, Schwarzenbacher H, Szczerbal I, Lohi H, Fries R, Taylor JF, Switonski M, Andersson M (2010) A novel mutation in the maternally imprinted PEG3 domain results in a loss of MIMT1 expression and causes abortions and stillbirths in cattle (Bos taurus). PLoS One 5:e15116–e15116

321. Hofstetter S, Seefried F, Häfliger IM, Jagannathan V, Leeb T, Drögemüller C (2019) A non-coding regulatory variant in the 5′-region of the MITF gene

is associated with white-spotted coat in Brown Swiss cattle. Anim Genet 50:27–32

322. Fontanesi L, Scotti E, Russo V (2012) Haplotype variability in the bovine MITF gene and association with piebaldism in Holstein and Simmental cattle breeds. Anim Genet 43:250–256

323. Wiedemar N, Drögemüller C (2014) A 19-Mb de novo deletion on BTA 22 including MITF leads to microphthalmia and the absence of pigmentation in a Holstein calf. Anim Genet 45:868–870

324. Philipp U, Lupp B, Mömke S, Stein V, Tipold A, Eule JC, Rehage J, Distl O (2011) A MITF mutation associated with a dominant white phenotype and bilateral deafness in German Fleckvieh cattle. PLoS One 6:e28857

325. Murgiano L, Jagannathan V, Piffer C, Diez-Prieto I, Bolcato M, Gentile A, Drögemüller C (2016) A frameshift mutation in MOCOS is associated with familial renal syndrome (xanthinuria) in Tyrolean Grey cattle. BMC Vet Res 12:1–10

326. Grobet L, Martin LJR, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Ménissier F, Massabanda J (1997) A deletion in the bovine myostatin gene causes the double–muscled phenotype in cattle. Nat Genet 17:71–74

327. Kambadur R, Sharma M, Smith TPL, Bass JJ (1997) Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. Genome Res 7:910–915

328. Grobet L, Poncelet D, Royo LJ, Brouwers B, Pirottin D, Michaux C, Ménissier F, Zanotti M, Dunner S, Georges M (1998) Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. Mamm genome 9:210–213

329. O'Rourke BA, Greenwood PL, Arthur PF, Goddard ME (2013) Inferring the recent ancestry of myostatin alleles affecting muscle mass in cattle. Anim Genet 44:86–90

330. Dierks C, Eder J, Glatzer S, Lehner S, Distl O (2014) A novel myostatin mutation in double-muscled German Gelbvieh. Anim Genet 46:91–92

331. Berry C, Thomas M, Langley B, Sharma M, Kambadur R (2002) Single cysteine to tyrosine transition inactivates the growth inhibitory function of Piedmontese myostatin. Am J Physiol Physiol 283:C135–C141

332. Denholm LJ, Martin LE, Teseling CF, Parnell PF, Beever JE (2014) Developmental Duplications (DD): 2. Mutation of the NHLRC2 gene causes neural tube defects in Angus cattle with multiple congenital malformation phenotypes that include axial and limb duplications, heteropagus conjoined twins, midbrain and forebrain malfo. Proc. XXVIII World Buiatrics Congr. Cairns Aust. Abstr. 100:

333. Owczarek-Lipska M, Plattet P, Zipperle L, Drögemüller C, Posthaus H, Dolf G, Braunschweig MH (2011) A nonsense mutation in the optic atrophy 3 gene (OPA3) causes dilated cardiomyopathy in Red Holstein cattle. Genomics 97:51–57

334. Simpson MA, Cook RW, Solanki P, Patton MA, Dennis JA, Crosby AH

(2009) A mutation in NFκB interacting protein 1 causes cardiomyopathy and woolly haircoat syndrome of Poll Hereford cattle. Anim Genet 40:42–46

335. Littlejohn MD, Henty KM, Tiplady K, Johnson T, Harland C, Lopdell T, Sherlock RG, Li W, Lukefahr SD, Shanks BC, Garrick DJ, Snell RG, Spelman RJ, Davis SR (2014) Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nat Commun 5:1–8

336. Boudreaux MK, Schmutz SM, French PS (2007) Calcium diacylglycerol guanine nucleotide exchange factor I (CalDAG-GEFI) gene mutations in a thrombopathic Simmental calf. Vet Pathol 44:932–935

337. Porto-Neto LR, Bickhart DM, Landaeta-Hernandez AJ, Utsunomiya YT, Pagan M, Jimenez E, Hansen PJ, Dikmen S, Schroeder SG, Kim E-S, Sun J, Crespo E, Amati N, Cole JB, Null DJ, Garcia JF, Reverter A, Barendse W, Sonstegard TS (2018) Convergent evolution of slick coat in cattle through truncation mutations in the prolactin receptor. Front Genet 9:57

338. Pausch H, Kölle S, Wurmser C, Schwarzenbacher H, Emmerling R, Jansen S, Trottmann M, Fuerst C, Götz K-U, Fries R (2014) A nonsense mutation in TMEM95 encoding a nondescript transmembrane protein causes idiopathic male subfertility in cattle. PLOS Genet 10:e1004044

339. Shanks RD, Dombrowski DB, Harpestad GW, Robinson JL (1984) Inheritance of UMP synthase in dairy cattle. J Hered 75:337–340

340. Schwenger B, Schöber S, Simon D (1993) DUMPS cattle carry a point mutation in the uridine monophosphate synthase gene. Genomics 16:241–244

341. Capitan A, Allais-Bonnet A, Pinton A, Marquant-Le Guienne B, Le Bourhis D, Grohs C, Bouet S, Clément L, Salas-Cortes L, Venot E, Chaffaux S, Weiss B, Delpeuch A, Noé G, Rossignol M-N, Barbey S, Dozias D, Cobo E, Barasc H, Auguste A, Pannetier M, Deloche M-C, Lhuilier E, Bouchez O, Esquerré D, Salin G, Klopp C, Donnadieu C, Chantry-Darmon C, Hayes H, Gallard Y, Ponsart C, Boichard D, Pailhoux E (2012) A 3.7 Mb deletion encompassing ZEB2 causes a novel polled and multisystemic syndrome in the progeny of a somatic mosaic bull. PLoS One 7:e49084

342. Gehrke LJ, Upadhyay M, Heidrich K, Kunz E, Klaus-Halla D, Weber F, Zerbe H, Seichter D, Graf A, Krebs S, Blum H, Capitan A, Thaller G, Medugorac I (2020) A de novo frameshift mutation in ZEB2 causes polledness, abnormal skull shape, small body stature and subfertility in Fleckvieh cattle. Sci Rep 10:1–14

343. HUGO Gene Nomenclature Committee (2021) HUGO Gene Nomenclature Committee at the European Bioinformatics Institute. https://www.genenames.org. Accessed 21 Nov 2021

344. Cole JB, VanRaden PM, Null DJ, Hutchison JL, Cooper TA, Hubbard SM (2018) Haplotype tests for recessive disorders that affect fertility and other traits. In: Anim. Improv. Progr. Res. Rep. https://www.aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html. Accessed 15 Apr 2021

345. McKusick-Nathans Institute of Genetic Medicine JHU Online Mendelian Inheritance in Man. In: 2021. https://www.omim.org. Accessed 31 Oct 2021

10 Declaration of Originality

Declaration of Originality

Last name, first name: Häfliger, Irene Monika

Matriculation number: 11-921-095

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations.

All data, tables, figures and text citations which have been reproduced from any other source, including the internet, have been explicitly acknowledged as such.

I am aware that in case of non-compliance, the Senate is entitled to withdraw the doctorate degree awarded to me on the basis of the present thesis, in accordance with the "Statut der Universität Bern (Universitätsstatut; UniSt)", Art. 69, of 7 June 2011.

Place, date

Willisau, 20.01.2022

Signature