Understanding genetic and phenotypic diversity in the speciation of *Petunia*: evolution of quantitative and qualitative traits

Inaugural dissertation
of the Faculty of Science,
University of Bern

presented by
Marta Binaghi
from Italy

Supervisor of the doctoral thesis
Prof. Dr. Cris Kuhlemeier
Institute of Plant Sciences

Understanding genetic and phenotypic diversity in the speciation of *Petunia*: evolution of quantitative and qualitative traits

Inaugural dissertation
of the Faculty of Science,
University of Bern

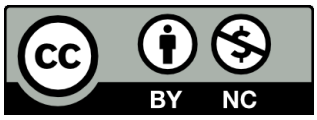
presented by
Marta Binaghi
from Italy

Supervisor of the doctoral thesis
Prof. Dr. Cris Kuhlemeier
Institute of Plant Sciences

Accepted by the Faculty of Science.

Bern, 13 December 2021

The Dean
Prof. Dr. Zoltan Balogh



Foreword

The experimental work presented here is the result of my doctoral project conducted at the Institute of Plant Sciences, University of Bern, between September 2017 and October 2021. I designed the projects and performed the experiments under the supervision of Prof. Dr. Cris Kuhlemeier. In some cases, colleagues have performed specific experiments and have helped me with experiment design and result analysis. The contributions to the two experimental chapters are stated in the author contribution paragraphs.

Acknowledgements

I would like to thank Prof. Dr. Cris Kuhlemeier for welcoming me in his group and for his supervision. I am really grateful for the opportunity to attend bioinformatics courses at the beginning of my PhD, and for the freedom that he granted me in the research work. Last but not least, his precious advice on writing has helped me a lot in these last few months.

The support I received by the research group is invaluable. From the first day I could enjoy a collaborative environment, and that allowed me to learn a lot. I have to thank every present and past member of the group and the office mates for their support. Ana, Andrea, Chaobin, Christopher, Diane, Doro, Fabrizio, Gina, Hagen, Jasmin, Jens, Korinna, Lea, Louis, Ludivine, Marc, Martina, Mathieu, Mathilde, Michel, Noëlle, Sarah D., Sarah R., Therese, Tracey, Tural, Vivien.

I also would like to spend a big “Thank you” to the software and methods developer that share their resources and material online, and provide detailed documentation. Some of them have personally helped me clarifying software use and identifying bugs.

Finally, I thank Kim for her precious advice on population genetics, and Prof. Loreta Brandao de Freitas for her collaboration on the projects involving wild accessions.

A PhD is a long, sometimes fun and sometimes tough journey. The people around me have majorly contributed to the fun parts, and have made the tough moments easier.

Summary

The diversity of flowering plants is connected with the recruitment of pollinators in the reproductive process. Some species of plants have specialised floral phenotypes that attract and reward one or few pollinators, and that allow efficient pollen exchange. The pollinator specificity of flowers can constitute a barrier to reproduction, by isolating plants pollinated by different animals. The group of traits (termed "pollination syndrome") that defines which pollinator will transfer pollen efficiently can thus define species isolation. In this manuscript I focus on the genetic basis of the changes in pollination syndrome traits in the *Petunia* genus.

P. axillaris and *P. exserta* have different pollinators, and their pollination syndrome traits differ in UV and visible colour, in scent production and in floral morphology. We investigated flower colour and morphology in individuals at two contact zones where the species live in sympatry and produce hybrid offspring. We used a genome-wide variant dataset to describe the population genetics of the individuals, and combined genotype information with phenotype data to estimate the genetic architecture of the traits and identify associated loci. The traits studied reveal different genetic architectures. Visible colour is associated with a polygenic basis, while UV colour with an oligogenic basis including less than 10 genes. Both the morphological traits (pistil and tube length) are predicted to have an oligogenic basis. We thus highlight that the classic association of discrete traits with a simple genetic basis, and continuous traits with a complex genetic basis is not always valid. The selection scan on these individuals identifies two broad regions experiencing positive selection. The regions do not include the speciation gene *MYB-FL*, responsible for the UV colour difference between the species. We suggest that this gene and the region associated with it are not under selection at the contact zones, maybe for a loss of importance of the trait in these populations. We

then suggest that the regions positively selected may be the result of selection pressure acting on the morphology phenotypes, or may be due to incompatibilities developed during the divergence of the species.

Scent production has changed several times during *Petunia* speciation. In particular, the shift from the bee pollination of *P. inflata* to the hawkmoth pollination of *P. axillaris* included an increase in the amount and number of volatiles produced. The biosynthesis of volatiles and the genetic changes that underlie scent production have been characterised in these species and have a simple genetic basis. On the other hand, the emission route of these volatiles has only been described in the research cultivar *P. hybrida* cv Mitchell. Scent production underwent important changes during *Petunia* evolution. We thus ask if these changes were accompanied by the adaptation of the scent emission system to sustain the new and increased production. We investigated the sequence conservation of the *P. hybrida* cv Mitchell transporter ABCG1 in the wild *Petunia* species, and we measured its expression in the strongly scented *P. axillaris*. *P. axillaris* and *P. hybrida* cv Mitchell share phenotypic similarities in their scent profiles as well as in other floral traits. *P. hybrida* cv Mitchell is derived from a series of crosses involving *P. axillaris* as a recurrent parent and a commercial variety descended from a species of the bee-pollinated short-tube clade (such as *P. inflata*). Consistently with its origin, the transcriptome of *P. hybrida* cv Mitchell is mostly constituted of *P. axillaris* transcript alleles. We thus hypothesised that the scent transporter ABCG1 function is conserved in *P. axillaris*. Unexpectedly we found that the sequence of the protein of *P. hybrida* is most similar to a species of the short-tube clade, suggesting that it was inherited from a modestly scented *Petunia*. Even more surprisingly, the *ABCG1* transporter is not expressed in the scented *P. axillaris*. We thus identify other *ABCG* subfamily members that could effect scent emission in this wild *Petunia* by analysing their expression profiles. We suggest that another gene may be responsible for scent emission in *P. axillaris*, or alternatively that *P. axillaris* does not need a transporter to emit scent. While the scent biosynthetic pathway is generally conserved between *P. hybrida* cv Mitchell and *P. axillaris*, scent emission took a different route.

Table of contents

Abbreviations

aLRT	Approximate likelihood-ratio test
AU	Absorbance unit
BSLMM	Bayesian sparse linear mixed model
CI	Confidence interval
cM	Centimorgans
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats - CRISPR associated
DE	Differential expression
DNA	Deoxyribonucleic acid
DNAseq	DNA sequencing
FC	Fold change
FDR	False discovery rate
FW	Fresh weight
GWAS	Genome-wide association study
KYA	Thousand years ago
LMM	Linear mixed model
LOD	Logarithm of the odds
MAF	Minor allele frequency
MYA	Million years ago
NBD	Nucleotide-binding domain
PC	Principal component
PCA	PC analysis
PCR	Polymerase chain reaction
PGE	Proportion of genetic variance
PVE	Proportion of variance in phenotypes
qRT-PCR	Quantitative RT-PCR
QTL	Quantitative trait locus
RNA	Ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
RT-PCR	Reverse-transcription PCR
SD	Standard deviation
TMD	Transmembrane domain
UV	Ultraviolet
VIGS	Virus-induced gene silencing

Genes, gene families, loci, enzymes

In the text, italicised names indicate the gene, roman names indicate the enzyme.

E.g. *ABCG1* is the gene that encodes the protein transporter ABCG1.

ABC	Adenosine triphosphate-binding cassette
ABCG1	ABC transporter subfamily G 1
ACTIN11	Actin-11
AN2	ANTHOCYANIN2
BPBT	Benzoyl-CoA:benzylalcohol/2-phenylethanol benzoyltransferase
BSMT	Benzoic acid/salicylic acid carboxyl methyltransferase
CNL	Trans-cinnamate:CoA ligase
DPL	DEEP PURPLE
EOBI	Emission of benzenoids I
EOBII	Emission of benzenoids II
LO2	Tomato locus of Style2.1
MYB-FL	R2R3-MYB transcription factor
ODO1	ODORANT1
PDR	Pleiotropic drug resistance
POPOVICH	C2H2 zinc-finger transcription factor
RAN1	Copper-transporting ATPase RAN1
WBC	White-brown complex
YUP	YELLOW UPPER

Chapter 1
Introduction

1.1 Speciation

> *The species is the principal unit of evolution.*
>
> *A sound understanding of the biological nature of species is fundamental to writing about evolution and indeed about almost any aspect of the philosophy of biology… I define biological species as 'groups of interbreeding natural populations that are reproductively (genetically) isolated from other such groups.'*
>
> *— Ernst Mayr as cited by Hey et al., 2005*

The definition of species has been intensely debated over the last century, but the biological species concept described by Mayr in the 1940s remains a valid framework for the study of speciation (Gao and Rieseberg, 2020). Mayr's biological species concept is based on the existence of reproductive isolation between diverging groups of individuals. Reproductive isolation barriers limit gene flow between incipient species, and thus constitute a central aspect of the speciation process. Understanding the evolution of reproductive isolation and the factors involved in it can therefore shed light on the process of speciation, on how a group of individuals diverges and becomes isolated as a species. Under Mayr's definition and for the purpose of this manuscript, the evolution of reproductive isolation is hence synonymous with speciation.

1.1.1 Reproductive isolation mechanisms

Reproductive isolation is achieved when one or more barriers prevent gene flow between groups of individuals, i.e. when alleles of one group cannot be transmitted to the other group or if after transmission they cannot persist. In plants, a difference in flowering period can constitute a reproductive barrier but also the inviability of hybrid seeds can limit the gene flow (e.g. Martin et al., 2007; Scopece et al., 2007). Reproductive barriers can act at different levels, and are generally classified into prezygotic and postzygotic, or prepollination and postpollination in the case of plants (Table 1.1). Prezygotic barriers prevent the meeting of the parental gametes by means of natural selection on one of the parents (an ecological factor), or by sexual selection. For example, organisms adapted to different habitats, or reproducing at different times will not exchange gametes. In the case of postzygotic isolation instead the gametes meet and a zygote is formed. The zygote in this case is sterile or inviable and cannot reproduce. The postzygotic barriers can be intrinsic such as developmental problems of the hybrid, or extrinsic such as an environmental factor hindering the survival or the reproduction of the hybrid.

The reproductive isolation between species can be determined by one or more barriers, and interactions between barriers have already been described (among others Coyne and Orr, 1997; Nosil et al., 2002; Ramsey et al., 2003; Widmer et al., 2009). The extent of the contribution of each barrier to the isolation of the incipient species can vary between organisms. Some species can be fully isolated simply due to unsuccessful fertilization (e.g. no hybrid seed formation in *Chamaecrista*, Costa et al., 2007), while others can reach full isolation by accumulating several barriers (e.g. ecogeography, pollinator specialisation, low germination and fertility of the hybrids in *Mimulus.* Ramsey et al., 2003). The relevance of each barrier in the speciation process, and in particular of the prezygotic versus postzygotic barriers, is generally calculated as its rate of evolution across different taxa (Coyne and Orr, 1989). Even though the studies comparing rates of evolution of different barriers are not numerous, many of them suggest that prezygotic isolation is more relevant, particularly in recently diverged species, although this pat-

tern is not universal (Matute and Cooper, 2021). Understanding how and why reproductive isolation barriers arise during speciation is a difficult challenge that includes: 1) the identification of each reproductive barrier and the precise quantification of their effects; 2) the identification of the molecular and genetic basis of the barriers; and 3) the understanding of the evolutionary forces that acted on each barrier (Lowry et al., 2008). In the next section I give an overview of the current knowledge on the genetic basis of reproductive isolation.

Table 1.1 Reproductive isolation barriers in plants. Barriers are listed in the order in which they act. Adapted from Baack et al., 2015.

Table adapted from that represented in Figure 1 of Baack et al., 2015.

https://doi.org/10.1111/nph.13424

1.1.2 The genetic basis of reproductive isolation

The natural variation that is found among living organisms is fundamental for the speciation process. Selection acts on the diversity between individuals and can determine the reproductive isolation leading to speciation. How individual diversity arises and provides the variation necessary for reproductive isolation is therefore a highly investigated question.

For most of the twentieth century, phenotypic diversity underlying evolution was thought to be due to the accumulation of a big number of small effect mutations (micromutationism). This view derived from Darwin's idea of an extremely gradual evolution, coupled with the advances in Mendelian genetics and supported by the infinitesimal model of Fisher (Orr, 2005). Towards the end of the century though, advances in the molecular techniques allowed the study of the genetic basis of phenotypic differences that characterise species. Several works revealed then that few loci of relatively large effect, termed "speciation genes", were responsible for phenotypic changes. Among others, the architectural differences between maize and its wild relative teosinte and the differences in larval morphology between *Drosophila* species (Doebley et al., 1995; Sucena and Stern, 2000). These results indicated that evolution could include few changes of large effect, and suggested that the number of required mutations was small compared to the "infinitesimal" number of Fisher's model.

The same technological advances allowed also the investigation of the genetic basis of specific reproductive isolation barriers. In principle, reproductive isolation can be affected by mutations at different scales: single nucleotide or small structural variants that affect protein structure or expression, gene duplication, chromosomal rearrangements and genome duplications. In plants, a number of studies have addressed the genetic basis of reproductive isolation barriers, revealing different architectures underlying them. Classical examples of barriers underlain by speciation genes include cytoplasmic male sterility (reviewed in Chase, 2007), hybrid necrosis (reviewed in Bomblies and Weigel, 2007), and flower colour

changes that result in pollinator shifts (Hoballah et al., 2007). Several loci are instead responsible for local adaptation to saline environments in *Heliantus paradoxus* (Lexer et al., 2003), and conspecific pollen precedence in *Mimulus* (Fishman et al., 2008). Structural variations such as chromosomal translocations or inversions could explain 50 % of the barrier to gene flow between *Heliantus annuus* and *H. petiolaris* (Rieseberg et al., 1999), and a single inversion partially contributes to the reproductive isolation between two ecotypes of *Mimulus guttatus* (Lowry and Willis, 2010). Thus different genetic architectures underlie these barriers, and a common mode for the genetic basis of speciation (admitting one would exist) has not yet been established.

In this context, another interesting aspect of the genetic basis of species difference is how sets of phenotypic traits can be maintained in the face of recombination. As a matter of fact, correlated phenotypic traits can be underlain by gene clusters that display low recombination; such regions are termed “supergenes” (reviewed recently in Gutiérrez-Valencia et al., 2021). Classic examples are the mimicry in the wing colour patterning of the butterfly *Heliconius numata* (Joron et al., 2011) and the *S* locus responsible for heterostyly in *Primula vulgaris* (Mather, 1950). In *Primula*, two morphs have different relative positions of style and anthers (heterostyly). This morphological difference promotes cross-breeding and is underlain by the genes in the *S* locus, a non-recombining region of 278 kb (Li et al., 2016). Supergenes can therefore maintain together the genetic differences that define species. The process of arising and maintenance of natural variation is still under scrutiny, as more and more studies highlight the genetic basis of different speciation traits.

For one thing it is clear: natural variation is determined by a diversity of genetic architectures, and only a wide investigation will allow students of the subject to find the patterns underlying it, and to confirm or disprove theoretical evolutionary models.

1.1.3 Pollinator-mediated speciation

Biotic pollination is the factor that has been more strongly associated with the high diversification between phyla during the evolution of land plants (Hernández-Hernández and Wiens, 2020). Changes in floral traits can determine the attraction of a new pollinator, and conversely changes in pollinator availability would require plant adaptation. Thus, pollinators can cause strong divergent selection pressure on flowering plants, and often constitute an important reproductive isolation barrier (Baack et al., 2015; Schemske and Bradshaw, 1999). As a matter of fact, different animals are attracted by different flowers based on visual and olfactory cues, and on the presence of a reward. Once the animal has approached the flower, the success of the pollination process depends on the morphology of the flower and in particular on the position of the reproductive organs, which has to ensure pollen transfer to and from the animal. Thus the adaptation of a flower to a specific pollinator includes changes in several phenotypic traits. The set of floral traits associated with a specific pollinator or pollinator group is termed "pollination syndrome" (Fenster et al., 2004). Typically, pollination syndromes include: 1) attractive traits such as colour (visible and non-visible by humans), scent, flower shape and size; 2) rewarding traits such as nectar volume, concentration and composition; and 3) efficiency traits such as position and length of the reproductive organs (Bradshaw et al., 1995).

A description of the genetic basis of pollinator adaptation is essential to understand the evolution of pollinator-mediated reproductive isolation. The interest in this topic is also fuelled by the notion that a shift in pollinator availability requires a relatively quick adaptation of the plant that includes several phenotypic traits. In addition, plants with intermediate phenotypes (i.e. when certain traits are adapted to the new pollinators but others are not) are expected to have lower fitness and thus fewer chances to reproduce. Such an evolutionary study system therefore contributes to the long-standing debate on evolutionary changes being due to many mutations of small effect or to major-effect mutations.

Over the last 30 years several workers described the genetic architecture of pollination syndrome traits, and identified loci responsible for phenotypic differences between them. Among these, Fishman and colleagues described a polygenic architecture characterised by small-effect quantitative trait loci (QTLs) in the mating system difference between *Mimulus guttatus* and *M. nasutus* (Fishman et al., 2002). Nakazato described several minor QTLs underlying the differences between *Ipomopsis guttata* and *tenuifolia*, and two QTLs of relatively large effect were suggested for the change in flower colour (Nakazato et al., 2013). Alongside these polygenic architectures, several genes of major effect were found to be sufficient for pollination syndrome differences (see for example the difference in scent volatiles in sexually deceptive orchids, Xu et al., 2012, and the colour difference in monkeyflowers, Yuan et al., 2013). Among these, only few studies tested the effect of the putative speciation loci on the pollinators.

In *Mimulus*, Bradshaw and Schemske were able to show that the *YUP* locus alleles from *M. lewisii* (*YUP*) and *M. cardinalis* (*yup*), responsible for carotenoid deposition, could attract the corresponding pollinator in near isogenic lines (Bradshaw and Schemske, 2003). In the same species, a loss of function mutation in the gene *OCIMENE SYNTHASE* could alter scent emission and result in decreased bumblebee visits to the flowers (Byers et al., 2014). In *Antirrhinum*, Shang and colleagues showed that the venation pattern caused by the gene *Venosa* could attract bumblebees as much as the original red morph, suggesting that the acquisition of the venation phenotype could constitute a quick adaptation to bumblebees (Shang et al., 2011). In *Petunia*, the transcription factor *ANTHOCYANIN2* (*AN2*) is responsible for colour difference between *P. axillaris* (hawkmoth-pollinated) and *P. integrifolia* (bee-pollinated, Quattrocchio et al., 1999). Hoballah and colleagues showed that bumblebees and hawkmoths preferred coloured and non-coloured flowers respectively, even when such difference was determined by the gene *AN2* alone in an otherwise identical genetic background (Hoballah et al., 2007). With a similar approach, Klahre and colleagues introgressed two loci responsible for scent production in the scentless *P. exserta* (hummingbird-pollinated), and conversely introgressed the *P. exserta* loci in *P. axillaris*, making it scentless. Pollina-

tor essays with hawkmoths showed that the animal preferred the scented plants, regardless of their colour (Klahre et al., 2011).

It is therefore clear that mutations of major effect can alter pollinator preference and therefore contribute to reproductive isolation between otherwise identical genetic backgrounds. This finding suggests that pollinator shifts could be initiated by one or few mutations and subsequently accumulate more small effect mutations as species diverge.

1.2 Wild petunias

The garden petunia (*P. hybrida*) has a long history as a genetic model species, which gained wide prominence with the discovery of the RNA interference mechanism (van der Krol et al., 1990; Napoli et al., 1990). In the last 20 years a considerable body of literature has accumulated on the wild relatives of *P. hybrida* too. In particular, the species of this genus have been studied under the aspect of population genetics and of the genetic basis of premating isolation barriers (Berardi et al., 2021; Bombarely et al., 2016; Caballero-Villalobos et al., 2021). In this section I present the genus *Petunia*, its different pollination syndromes, and a naturally occurring hybrid population.

1.2.1 The genus

The genus *Petunia* is endemic to temperate and subtropical regions of South America including Argentina, Bolivia, Brazil, Paraguay and Uruguay. The centre of diversity is located in southern Brazil, where some species can be widely distributed (e.g. *P. axillaris*) while others are endemic to specific and restricted environments (e.g. *P. exserta,* Stehmann et al., 2009). Interspecific crosses between *Petunia* species can produce viable offspring, indicating that reproductive barriers are mainly prezygotic (Ando et al., 1998, 2001; Kokubun et al., 2006; Watanabe et al., 2001). The geographic distribution of the species plays a substantial role in maintaining them isolated (Dell'Olivo et al., 2011). But when geographic barriers are not present, pollinator preferences for specific floral traits can still limit gene flow (Dell'Olivo and Kuhlemeier, 2013).

The flowers of the *Petunia* species present phenotypic differences in terms of colour and morphology. The length of the corolla tube in particular discriminates between two major clades: the short-tube and the long-tube clade. These clades have diverged recently, and while the short tube-clade includes species with uniform floral phenotypes, the long-tube clade provides an astonishing floral diversity (Reck-Kortmann et al., 2014; Särkinen et al., 2013; Stehmann et al., 2009). The recent divergence (2.85-1.3 million years ago, MYA. Särkinen et al., 2013) of the

long-tube clade and its phenotypic diversity have made it a clade of great interest for evolutionary, population genetic, and molecular studies; among others Ando et al., 2001; Berardi et al., 2021; Dell'Olivo et al., 2011; Hoballah et al., 2007; Lorenz-Lemke et al., 2006; Quattrocchio et al., 1999; Reck-Kortmann et al., 2017; Rodrigues et al., 2018a; Turchetto et al., 2015a.

1.2.2 Pollination syndromes in the genus *Petunia*

Most flowering plants rely on animals for sexual reproduction, and changes in the plant-pollinator interaction can lead to reproductive isolation and speciation (Ollerton et al., 2011). Such pollinator shifts are therefore widely studied in order to better understand the processes that underlie species evolution (Hodges and Arnold, 1994; Kay and Sargent, 2009; Martin et al., 2008; Whittall and Hodges, 2007).

The genus *Petunia* includes different pollination syndromes and the selection for different pollinators is considered to be an important force driving its floral diversification (Fregonezi et al., 2013). As a matter of fact, shifts in attractive traits influence pollinator visitation in *P. axillaris* and *P. inflata* (Dell'Olivo et al., 2011). The multiple pollinator shifts in this young clade have brought the attention to how evolution could attain multiple times such conspicuous phenotypic changes, whose intermediate steps have likely low fitness, in a relatively short period of time.

In *Petunia*, the short tube clade displays a bee-pollination syndrome, while in the long-tube clade we find three different pollinators (Figure 1.1). *P. inflata* (a representative of the short-tube clade) has purple flowers with a short and wide corolla tube, where bees can access both nectar and pollen (Ando et al., 2005; Dell'Olivo and Kuhlemeier, 2013; Gübitz et al., 2009). In the long-tube clade, however, we find *P. axillaris*, *P. exserta* and *P. secreta*, the three species presenting a long and narrow corolla tube. *P. axillaris*, with its white and UV-absorbent petals, and its rich floral bouquet of volatiles emitted at dusk, displays a hawkmoth pollination syndrome (Ando et al., 2001; Hermann et al., 2013; Hoballah et al., 2005). The

ability to distinguish UV-absorbent patterns in low light conditions and to perceive scent volatile compounds with their antennas make hawkmoths a perfect fit for this flower. *P. exserta*, with red petals presenting protruding reproductive organs and no scent emission, attracts hummingbirds, which are usually associated with red colour and are not able to perceive volatiles (Hermann et al., 2013, 2015; Lorenz-Lemke et al., 2006). *P. secreta* has purple flowers, placing it in the bee pollination syndrome, but maintains the elongated corolla which differentiates it from *P. inflata* (Stehmann and Semir, 2005; Stehmann et al., 2009). The narrow and long corolla of *P. secreta* is in fact inaccessible to bees, which cannot collect the nectar from these flowers. Nonetheless the flowers are effectively pollinated by pollen-collecting bees in the field, suggesting that the seemingly incompatible traits of these flowers might be ancestral characteristics not yet lost (Rodrigues et al., 2018b).

The current knowledge on the molecular basis of pollination syndromes in the *Petunia* long-tube clade is presented in Section 1.3.1.

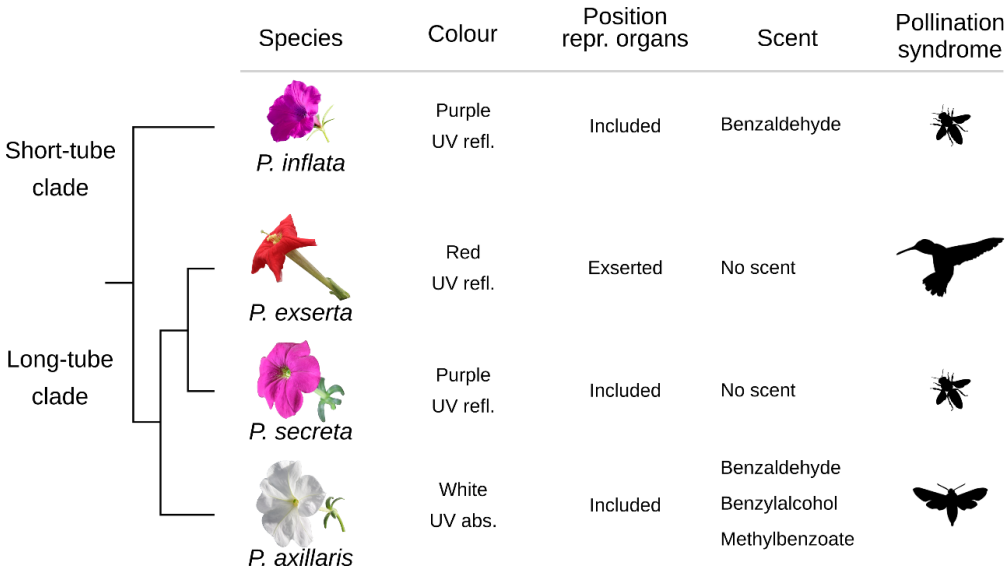


Figure 1.1 Pollination syndromes in the genus *Petunia*. The figure displays a cladogram of some species of the short and long-tube clade of *Petunia*, according to Esfeld et al., 2018 and based on 399,673 nucleotides. The colour in the visible and UV spectra is indicated. "Refl." reflective, "abs." absorbent. The position of the reproductive organs relative to the tube opening, and the main scent compounds produced by the flowers are indicated. The pollination syndromes are depicted by the animal silhouettes (bee, hummingbird and hawkmoth).

1.2.3 The *P. axillaris* × *P. exserta* hybrids

Reproductive barriers between Petunia *species*

A fundamental reproductive barrier between *Petunia* species is geography (Dell'Olivo et al., 2011). Different species and subspecies often occupy restricted territories and don't overlap (Stehmann et al., 2009). In addition the pollen dispersal is extremely limited in space in *P. axillaris*, and seed dispersal happens by autochory, leaving the seeds fall to the ground when the capsule opens, or shattering them at short distance from the mother plant (Stehmann et al., 2009; Turchetto et al., 2015b). The importance of the geographic barrier between these species is apparent also given the fact that different species can produce fertile offspring when artificially crossed, indicating a minor importance (in most crosses) of post-mating barriers (Watanabe et al., 2001). In the case of species that share a geographic region, it is therefore important for the maintenance of reproductive isolation that other barriers exist. One such barrier is described in the previous section: pollinator specificity. In the south of Brazil, the co-occurrence of different *Petunia*

species from the long-tube clade offers the opportunity to investigate the maintenance of species delimitation between interfertile species that display different pollination syndromes. This is the case of the contact zones between *P. axillaris* and *P. exserta*.

P. axillaris *and* P. exserta *habitats*

The geographic distribution of *P. axillaris* covers a vast territory; plants of this species occur in Argentina, Bolivia, South of Brazil, Paraguay and Uruguay (Lorenz-Lemke et al., 2006; Stehmann et al., 2009). They are often found in open habitats exposed to direct sunlight, in rocky outcrops, and also in disturbed locations (e.g. along the roadsides, Dell'Olivo et al., 2011; Lorenz-Lemke et al., 2006). The other members of the long-tube clade instead occupy smaller geographic regions (Stehmann et al., 2009). *P. exserta* in particular is endemic of the Serra do Sudeste region, in the south of Brazil (Lorenz-Lemke et al., 2006). *P. exserta* plants can only be found in small groups of individuals, on sandstone towers typical of the region. On the towers, the plants are located in shallow caves, where they are not exposed to the direct sunlight due to the southeast direction of the shelter opening, and they grow on a shallow soil (Lorenz-Lemke et al., 2006, Martina Lüthi personal communication). The lack of *P. exserta* individuals outside of the shelters does not seem to be related to the need for a shadowy environment for the germination (unpublished results mentioned in Lorenz-Lemke et al., 2006).

P. axillaris × P. exserta *hybrids*

Even though *P. axillaris* and *P. exserta* have different microhabitats (*P. axillaris* preferring open and sunny spots, while *P. exserta* grows in shady and rather unhospitable places), these two species have been found growing in close proximity (less than ten metres) at two sites in the Guaritas region (from here on, indicated as "contact zones", Lorenz-Lemke et al., 2006. Figure 1.2). Here, some *P. axillaris*-looking plants grow within a few metres of the *P. exserta* shelters, while *P. exserta*-looking individuals were never found outside of their shelters (Lorenz-Lemke et al., 2006, Figure 1.2).

In these two contact zones, individuals with intermediate floral phenotypes are found inside the shelters. These plants have petals that vary in colour from bright pink to almost white, and were described as having weakly exerted reproductive organs (Lorenz-Lemke et al., 2006; Schnitzler et al., 2020; Martina Lüthi personal communication). The two species produce fertile offspring when artificially crossed and the hybrid-looking individuals where classified as hybrids by several molecular studies, described below (Watanabe et al., 2001).

P. axillaris *and* P. exserta *"molecular" history*

The long-tube clade of *Petunia* diverged recently, developing quickly its different pollination syndromes (Särkinen et al., 2013). In this clade, *P. axillaris* and *P. exserta* share genetic polymorphisms and their divergence time is estimated at 310 thousand years ago (KYA. 99 % confidence interval (CI): 251-318 KYA. Caballero-Villalobos et al., 2021; Segatto et al., 2014; Turchetto et al., 2015a). The latest study by Caballero-Villalobos and colleagues on the molecular evolution history of these species reports a period of divergence followed by a secondary contact 0.92 KYA (99 % CI: 0.25-81.1 KYA), and ongoing hybridization with migration rate higher in the direction from *P. axillaris* to *P. exserta* (Caballero-Villalobos et al., 2021). The directionality could be due to the reduced reproductive success of *P. exserta* as a pollen donor shown in artificial crossings by Watanabe et al. (2001) or to the major abundance of *P. axillaris* and its higher diversity, compared to the inbreeding recorded in the *P. exserta* populations (Caballero-Villalobos et al., 2021; Watanabe et al., 2001).

In spite of the ongoing hybridisation between these two species, individuals displaying the pure species phenotype can be found in the hybrid zones (Teixeira et al., 2019, 2020). This observation supports a strong influence of pollinators on the reproductive isolation of these two species.

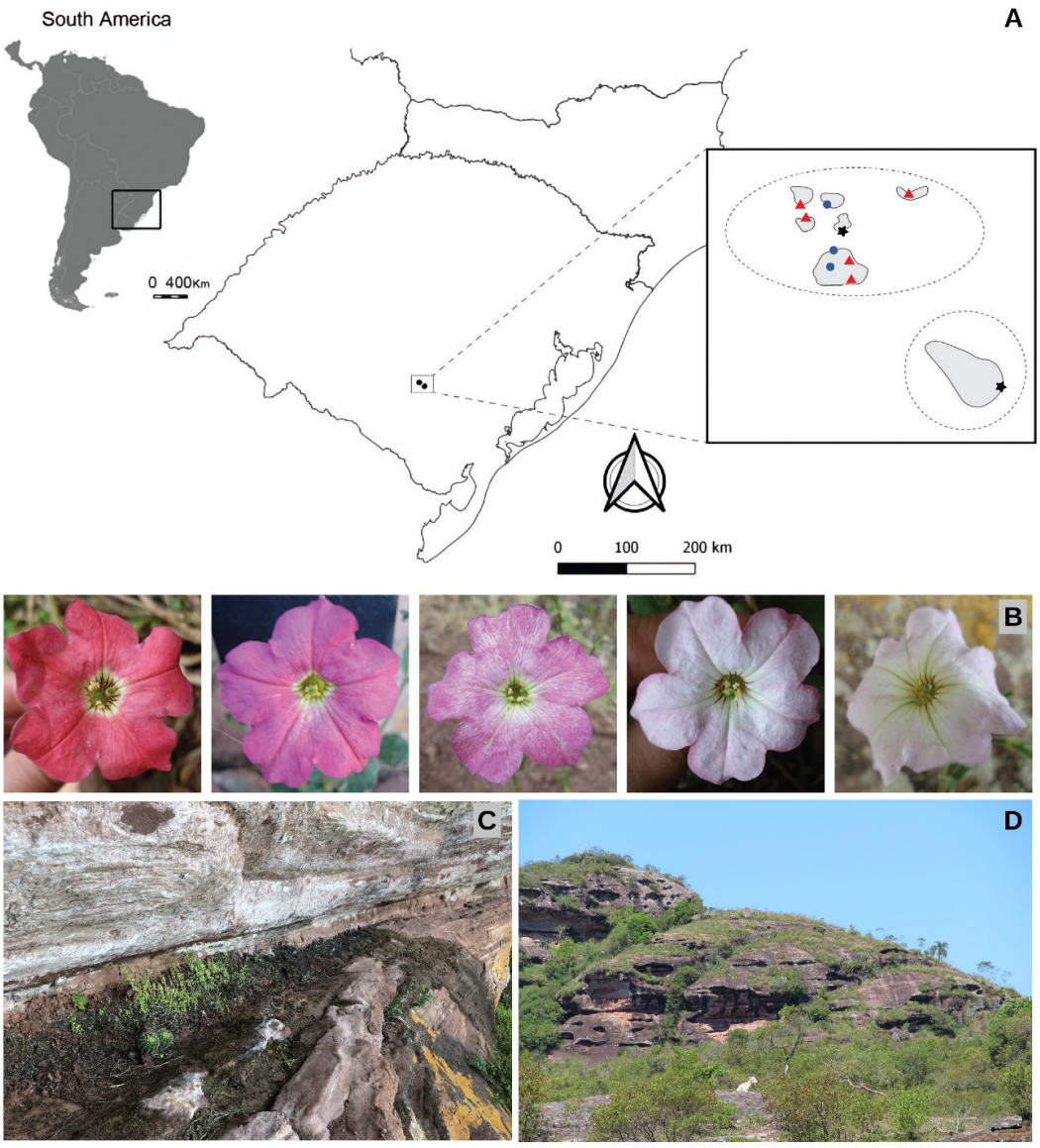


Figure 1.2 *P. axillaris* × *P. exserta* hybrids. A) Geographic location of the contact zones (black stars in the inset), and the nearby *P. axillaris* (blue dots) and *P. exserta* (red triangles) populations. B) Intermediately coloured flowers from the contact zones in frontal view. Panel A and B reproduced from Schnitzler et al., 2020. C and D habitats of *P. exserta* and *P. axillaris* respectively. Photos in C and D by Martina Lüthi.

1.3 *Petunia* as a model for evolutionary biology

The use of *Petunia* as a model plant extends to various research fields like molecular biology, genetics, comparative developmental biology, evolutionary biology and biochemistry of floral pigments and volatiles (Berardi et al., 2021; Faraco et al., 2017; van der Krol et al., 1990; Morel et al., 2018; Napoli et al., 1990; Nouri et al., 2021). The broad diffusion of Petunia as a research model was initially motivated by its transformation capacity and its active transposons that provided a source of tractable mutations, besides other useful laboratory characteristics such as a short life cycle and easy propagation (Vandenbussche et al., 2016). More recently, the *Petunia* resources have grown even more with the sequencing of the genomes of *P. axillaris* and *P. inflata* and with the development of CRISPR-Cas9 protocols for efficient transformation (Bombarely et al., 2016; Zhang et al., 2021). Moreover the community that gathers around this model plant is to be mentioned for its collaborative and friendly environment where advances in the research are shared regularly at international meetings, materials and data are exchanged easily between groups and collective efforts are made for the benefit of the whole community (such as the sequencing of the reference genomes).

In this section, I introduce the current knowledge on the evolutionary biology of *Petunia*.

1.3.1 The speciation genes of *Petunia*

In the last fifteen years, the molecular basis of evolution in the genus *Petunia* has revealed a number of simple mutations with large phenotypic effect at the base of the pollination syndrome shifts.

From bees to moths

The shift from the bee-pollinated common ancestor of *P. inflata* and *P. axillaris* to the hawkmoth-pollinated *P. axillaris* included changes of multiple phenotypic traits associated to pollinator attraction. The colour of the flowers turned from purple to white, the petals became UV-absorbent to allow moths to see the flowers

when the light is low, and the emission of volatiles increased (Dell'Olivo and Kuhlemeier, 2013). The genetic changes underlying these phenotypic differences have been identified (Table 1.2).

Table 1.2 Speciation genes in the *P. inflata* - *P. axillaris* shift

Trait	*P. inflata*	*P. axillaris*	Locus	Polymorphism	Reference
Visible colour	Purple	White	*AN2*[a]	Loss-of-function	Hoballah et al., 2007; Quattrocchio et al., 1999
UV colour	Reflective	Absorbent	*MYB-FL*[b]	Cis-upregulation	Sheehan et al., 2016
Volatile compounds	-	New benzoate compounds	*BPBT*, *BSMT*[c]	Activation of the enzymes	Amrad et al., 2016
Volatile quantity	Low	High		Activation of a TF (hypothesised)	Amrad et al., 2016

[a] *ANTHOCYANIN2*
[b] *MYB* transcription factor
[c] *Benzoyl-CoA:benzylalcohol/2-phenylethanol benzoyltransferase*, *benzoic acid/salicylic acid carboxyl methyltransferase*

From moths to hummingbirds

The shift to hummingbird pollination in *P. exserta* included the gain of red colour in the petals, the loss of UV absorbance, and the loss of scent production (Amrad et al., 2016; Berardi et al., 2021; Hermann et al., 2015; Klahre et al., 2011; Sheehan et al., 2016). Besides these attractive traits, a potential pollination efficiency trait was also changed. As a matter of fact, the position of the reproductive organs (style and stamens) changed from being included in the corolla tube to being exserted above the flower opening (Hermann et al., 2013). The genetic changes underlying some of these phenotypic differences have been identified, and in some cases QTLs responsible for the phenotype have been mapped (Table 1.3).

Table 1.3 Speciation genes in the *P. axillaris* - *P. exserta* shift

Trait	*P. axillaris*	*P. exserta*	Locus	Polymorphism	Reference
Visible colour	White	Red	*DPL*[a] + others	Upregulation of *DPL* and DE[b] of anthocyanin-transformation	Berardi et al., 2021
UV colour	Absorbent	Reflective	*MYB-FL*	1 bp deletion	Sheehan et al., 2016
Volatile production	High	Absent	*CNL*[c]	Null mutations in coding sequence	Amrad et al., 2016
Volatile production	High	Absent	*ODO1*[d]	Downregulation	Amrad et al., 2016
Repr. org. position[e]	Included	Exserted	Few QTLs of major effect		Hermann et al., 2013

[a] *DEEP PURPLE*
[b] Differential expression
[c] *Cinnamate-CoA ligase*
[d] *ODORANT1*
[e] Reproductive organ position

From moths to bees

The shift to bee-pollination in the species *P. secreta* included the re-gain of purple colour (Rodrigues et al., 2018b). Unexpectedly, the resurrection of colour in the petals is due to a 2 bp deletion that restores the reading frame of *AN2*, indicating that this *MYB* transcription factor is a hotspot for evolutionary changes in colour (Esfeld et al., 2018).

In many cases, shifts in traits involved in pollinator attraction have therefore shown to have a simple genetic basis. The genetics of the morphological differences between *P. axillaris* and *P. exserta* instead has proven harder to clarify. Hermann and colleagues investigated the genetic basis of the difference in reproductive organ lengths between the two species (Hermann et al., 2015). Their study identified a major QTL on chromosome 2 that reproduces the *P. axillaris* pistil length when introduced in the *P. exserta* background (Figure 1.3). In a *P. axillaris* background instead, the restoration of pistil length was only achieved by introgressing the chromosome 2 QTL and an additional QTL on chromosome 5. A

more recent morphometric analysis coupled with time-course RNAseq on *P. axillaris* and *P. exserta* identified several candidate genes for the difference in style length (Yarahmadov et al., 2020). Two of them, *EOBI* and *EOBII* (*emission of benzenoids I and II*) were further confirmed to affect style length by virus-induced gene silencing, suggesting that the phenylpropanoid/flavonoid pathway might be involved not only in floral colour and volatile changes, but also in morphological differences in the long-tube clade of *Petunia*.

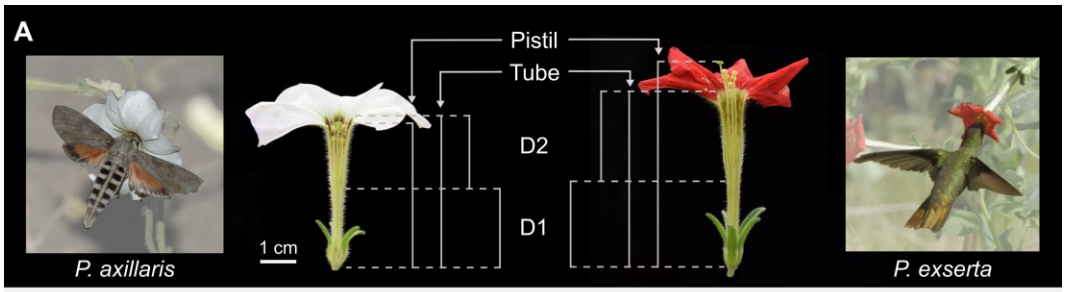


Figure 1.3 Morphological differences and their genetic basis between *P. axillaris* and *P. exserta*. A) Longitudinal section of *P. axillaris* and *P. exserta* showing differences in the length of floral organs. Photographs in the insets show the behaviour of the pollinators on the respective species. B, C and D, QTL analysis for reproductive organ and floral morphology (adapted from Hermann et al., 2015). The *y* axes indicate the LOD scores for each trait. Horizontal lines the significance threshold ($P < 0.01$) as determined by permutation analysis. The *x* axes delimit the seven linkage groups; mapping distances are given in centimorgans. pl pistil length.

Thus in *Petunia*, a few mutations have been shown to have major phenotypic effects on floral colour and scent evolution. The position of the reproductive organs has also a relatively simple genetic basis constituted by few QTLs, and recent results point to the direction of a few transcription factors involved in its regulation. The long-tube clade of *Petunia* indicates that evolution of different pollination syndromes involves few major effect genes.

1.3.2 The supergene

How species can inherit sets of traits and maintain them in the face of gene flow is a long-standing question. In *Petunia*, the pollination syndrome traits are controlled by different genes in different pathways. Nonetheless, a single QTL on chromosome 2 is responsible for colour, UV absorption and pistil length variation between *P. axillaris* and *P. exserta*, while the addition of a QTL from chromosome 7 restores scent production (Hermann et al., 2013). Hermann and colleagues observed that recombination in the chromosome 2 QTL was suppressed, as they could only find one recombination event that separated scent from the other traits in a progeny of 504 plants, indicating a distance of 0.1 cM between the locus for scent production and for the other traits on chromosome 2. Further experiments could identify more recombinants along the chromosome 2 locus, therefore excluding the role of a pleiotropic master regulator in controlling the set of phenotypes. Thus, the genetic basis of these pollination syndromes is located in a single genomic region that is maintained by suppression of recombination. This observation indicates the possible presence of a cluster of genes tightly linked, as in the case of a supergene. The existence of a supergene that controls the pollination syndrome traits in *P. axillaris* and *P. exserta* would limit the presence of unfit recombinant phenotypes and possibly contribute to species divergence.

With the advent of chromosome-length genome assemblies, this speciation gene cluster is being studied in more detail. The cause of the reduced recombination rate in the locus and its conservation in other interspecific crosses is addressed by the work of Tracey Tenreira (personal communication).

1.4 Subject of the thesis

Humans have been fascinated by the diversity of Nature for thousands of years. The diversity of the living organisms is the result of evolutionary processes that shaped them into complex entities, adapted to survive to different conditions. The changes that underlie evolution are the focus of this manuscript. With this work I hope to provide additional information on the evolutionary processes that act at the genetic level.

The work is centred on the model organism *Petunia*, whose wild species constitute a convenient system for the study of speciation. The *Petunia* genus is introduced in the current chapter (Chapter 1) together with the concept of speciation and its connection with pollination syndromes. The experimental work is divided in two experimental chapters (Chapter 2 and 3).

Chapter 2 investigates the importance of the pollination syndrome traits when two differently pollinated species, *P. axillaris* and *P. exserta*, come into secondary contact. More precisely, I ask how the speciation loci identified in interspecific crosses segregate in natural populations, and if selection can be detected on the loci associated to the traits. We use plant material from two contact zones of *P. axillaris* and *P. exserta* to obtain phenotype and genotype data. Whole-genome sequencing is used to obtain a set of variants that constitutes the core of the experiment. This set is used to observe genetic population structure, and is combined with phenotype data to perform a genome-wide association study (GWAS). The genotype data is finally scanned for selection signatures.

Chapter 3 focuses on a single pollination syndrome trait (scent) which has experienced several changes during pollinator shifts across the genus *Petunia*. The changes in this trait are due to specific mutations affecting volatile production; these changes in production result into different quantity and quality of the volatiles emitted in the environment. In general, the emission of a volatile compound can happen via active transport or diffusion. Recently, the emission of scent in the research cultivar *P. hybrida* was found to rely on active transport, and a reduction

in the transporter expression resulted in accumulations of the volatiles inside the cells (Adebesin et al., 2017). In *P. hybrida*, the active transport mediated by *ABCG1* is therefore necessary to guarantee scent emission. In the present work, we ask if the role of *ABCG1* is conserved in the wild *Petunia* species, and how the transporter is associated to the changes in the species' scent profiles that accompanied the pollinator shifts of this genus. We use sequence alignments to identify the homologues of *ABCG1* in the wild species, and RNAseq and qRT-PCR to quantify their expression. Finally we use sequence similarity and motif searches to characterise the *ABCG* subfamily in *P. axillaris*. The expression patterns of the subfamily members are analysed to identify other potential scent transporters.

The last chapter (Chapter 4) presents a general discussion of the results of the experimental sections and puts these results in a broader context.

1.5 References

Adebesin, F., Widhalm, J.R., Boachon, B., Lefèvre, F., Pierman, B., Lynch, J.H., Alam, I., Junqueira, B., Benke, R., Ray, S., et al. (2017). Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. Science *356*, 1386–1388.

Amrad, A., Moser, M., Mandel, T., de Vries, M., Schuurink, R.C., Freitas, L., and Kuhlemeier, C. (2016). Gain and loss of floral scent production through changes in structural genes during pollinator-mediated speciation. Current Biology *26*, 3303–3312.

Ando, T., Tsukamoto, T., Akiba, N., Kokubun, H., Watanabe, H., Ueda, Y., and Marchesi, E. (1998). Differentiation in the degree of self-incompatibility in *Petunia axillaris* (Solanaceae) occurring in Uruguay. Acta Phytotaxonomica et Geobotanica *49*, 37–47.

Ando, T., Nomura, M., Tsukahara, J., Watanabe, H., Kokubun, H., Tsukamoto, T., Hashimoto, G., Marchesi, E., and Kitching, I.J. (2001). Reproductive isolation in a native population of *Petunia sensu* Jussieu (Solanaceae). Annals of Botany *88*, 403–413.

Ando, T., Ishikawa, N., Watanabe, H., Kokubun, H., Yanagisawa, Y., Hashimoto, G., Marchesi, E., and Suárez, E. (2005). A morphological study of the *Petunia integrifolia* complex (Solanaceae). Annals of Botany *96*, 887–900.

Baack, E., Melo, M.C., Rieseberg, L.H., and Ortiz-Barrientos, D. (2015). The origins of reproductive isolation in plants. New Phytologist *207*, 968–984.

Berardi, A.E., Esfeld, K., Jäggi, L., Mandel, T., Cannarozzi, G.M., and Kuhlemeier, C. (2021). Complex evolution of novel red floral color in *Petunia*. Plant Cell *33*, 2273–2295.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C.S., Bliek, M., Boersma, M.R., Borghi, L., Bruggmann, R., et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants *2*, 16074.

Bomblies, K., and Weigel, D. (2007). Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. Nature Reviews Genetics *8*, 382–393.

Bradshaw, H.D., and Schemske, D.W. (2003). Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature *426*, 176–178.

Bradshaw, H.D., Wilbert, S.M., Otto, K.G., and Schemske, D.W. (1995). Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). Nature *376*, 762–765.

Byers, K.J.R.P., Vela, J.P., Peng, F., Riffell, J.A., and Bradshaw, H.D. (2014). Floral volatile alleles can contribute to pollinator-mediated reproductive isolation in monkeyflowers (*Mimulus*). The Plant Journal *80*, 1031–1042.

Caballero-Villalobos, L., Silva-Arias, G.A., Turchetto, C., Giudicelli, G.C., Petzold, E., Bombarely, A., and Freitas, L.B. (2021). Neutral and adaptive genomic variation in hybrid zones of two ecologically diverged *Petunia* species (Solanaceae). Botanical Journal of the Linnean Society *196*, 100–122.

Chase, C.D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. Trends in Genetics *23*, 81–90.

Costa, C.B.N., Lambert, S.M., Borba, E.L., and de Queiroz, L.P. (2007). Post-zygotic reproductive isolation between sympatric taxa in the *Chamaecrista desvauxii* Complex (Leguminosae–Caesalpinioideae). Annals of Botany *99*, 625–635.

Coyne, J.A., and Orr, H.A. (1989). Patterns of speciation in *Drosophila*. Evolution *43*, 362–381.

Coyne, J.A., and Orr, H.A. (1997). "Patterns of speciation in *Drosophila*" revisited. Evolution *51*, 295–303.

Dell'Olivo, A., and Kuhlemeier, C. (2013). Asymmetric effects of loss and gain of a floral trait on pollinator preference. Evolution *67*, 3023–3031.

Dell'Olivo, A., Hoballah, M.E., Gübitz, T., and Kuhlemeier, C. (2011). Isolation barriers between *Petunia axillaris* and *Petunia integrifolia* (Solanaceae). Evolution *65*, 1979–1991.

Doebley, J., Stec, A., and Gustus, C. (1995). Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics *141*, 333–346.

Esfeld, K., Berardi, A.E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. (2018). Pseudogenization and resurrection of a speciation gene. Current Biology *28*, 3776-3786.e7.

Faraco, M., Li, Y., Li, S., Spelt, C., Di Sansebastiano, G.P., Reale, L., Ferranti, F., Verweij, W., Koes, R., and Quattrocchio, F.M. (2017). A tonoplast P3B-ATPase mediates fusion of two types of vacuoles in petal cells. Cell Reports *19*, 2413–2422.

Fenster, C.B., Armbruster, W.S., Wilson, P., Dudash, M.R., and Thomson, J.D. (2004). Pollination syndromes and floral specialization. Annu. Rev. Ecol. Evol. Syst. *35*, 375–403.

Fishman, L., Kelly, A.J., and Willis, J.H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. Evolution *56*, 2138–2155.

Fishman, L., Aagaard, J., and Tuthill, J.C. (2008). Toward the evolutionary genomics of gametophytic divergence: patterns of transmission ratio distortion in monkeyflower (*Mimulus*) hybrids reveal a complex genetic basis for conspecific pollen precedence. Evolution *62*, 2958–2970.

Fregonezi, J.N., Turchetto, C., Bonatto, S.L., and Freitas, L.B. (2013). Biogeographical history and diversification of *Petunia* and *Calibrachoa* (Solanaceae) in the Neotropical Pampas grassland. Botanical Journal of the Linnean Society *171*, 140–153.

Gao, L., and Rieseberg, L.H. (2020). While neither universally applicable nor practical operationally, the biological species concept continues to offer a compelling framework for studying species and speciation. National Science Review *7*, 1398–1400.

Gübitz, T., Hoballah, M.E., Dell'Olivo, A., and Kuhlemeier, C. (2009). Petunia as a model system for the genetics and evolution of pollination syndromes. In Petunia, T. Gerats, and J. Strommer, eds. (New York, NY: Springer New York), pp. 29–50.

Gutiérrez-Valencia, J., Hughes, P.W., Berdan, E.L., and Slotte, T. (2021). The genomic architecture and evolutionary fates of supergenes. Genome Biology and Evolution *13*.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T., and Kuhlemeier, C. (2013). Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. Current Biology *23*, 873–877.

Hermann, K., Klahre, U., Venail, J., Brandenburg, A., and Kuhlemeier, C. (2015). The genetics of reproductive organ morphology in two *Petunia* species with contrasting pollination syndromes. Planta *241*, 1241–1254.

Hernández-Hernández, T., and Wiens, J.J. (2020). Why Are There So Many Flowering Plants? A Multiscale Analysis of Plant Diversification. The American Naturalist *195*, 948–963.

Hey, J., Fitch, W.M., and Ayala, F.J. (2005). Systematics and the origin of species: an introduction. PNAS *102*, 6515–6519.

Hoballah, M.E., Stuurman, J., Turlings, T.C.J., Guerin, P.M., Connétable, S., and Kuhlemeier, C. (2005). The composition and timing of flower odour emission by wild *Petunia axillaris* coincide with the antennal perception and nocturnal activity of the pollinator *Manduca sexta*. Planta *222*, 141–150.

Hoballah, M.E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., and Kuhlemeier, C. (2007). Single gene–mediated shift in pollinator attraction in *Petunia*. The Plant Cell *19*, 779–790.

Hodges, S.A., and Arnold, M.L. (1994). Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. PNAS *91*, 2493–2496.

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., et al. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature *477*, 203–206.

Kay, K.M., and Sargent, R.D. (2009). The role of animal pollination in plant speciation: integrating ecology, geography, and genetics. Annu. Rev. Ecol. Evol. Syst. *40*, 637–656.

Klahre, U., Gurba, A., Hermann, K., Saxenhofer, M., Bossolini, E., Guerin, P.M., and Kuhlemeier, C. (2011). Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. Current Biology *21*, 730–739.

Kokubun, H., Nakano, M., Tsukamoto, T., Watanabe, H., Hashimoto, G., Marchesi, E., Bullrich, L., Basualdo, I.L., Kao, T., and Ando, T. (2006). Distribution of self-compatible and self-incompatible populations of *Petunia axillaris* (Solanaceae) outside Uruguay. J Plant Res *119*, 419.

van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in *Petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291–299.

Lexer, C., Welch, M.E., Durphy, J.L., and Rieseberg, L.H. (2003). Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. Molecular Ecology *12*, 1225–1235.

Li, J., Cocker, J.M., Wright, J., Webster, M.A., McMullan, M., Dyer, S., Swarbreck, D., Caccamo, M., Oosterhout, C. van, and Gilmartin, P.M. (2016). Genetic architecture and evolution of the *S* locus supergene in *Primula vulgaris*. Nature Plants *2*, 1–7.

Lorenz-Lemke, A.P., Mäder, G., Muschner, V.C., Stehmann, J.R., Bonatto, S.L., Salzano, F.M., and Freitas, L.B. (2006). Diversity and natural hybridization in a

highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. Molecular Ecology *15*, 4487–4497.

Lowry, D.B., and Willis, J.H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLOS Biology *8*, e1000500.

Lowry, D.B., Modliszewski, J.L., Wright, K.M., Wu, C.A., and Willis, J.H. (2008). The strength and genetic basis of reproductive isolating barriers in flowering plants. Philosophical Transactions of the Royal Society B: Biological Sciences *363*, 3009–3021.

Martin, N.H., Bouck, A.C., and Arnold, M.L. (2007). The genetic architecture of reproductive isolation in Louisiana irises: flowering phenology. Genetics *175*, 1803–1812.

Martin, N.H., Sapir, Y., and Arnold, M.L. (2008). The genetic architecture of reproductive isolation in Louisiana irises: pollination syndromes and pollinator preferences. Evolution *62*, 740–752.

Mather, K. (1950). The genetical architecture of heterostyly in *Primula sinensis*. Evolution *4*, 340–352.

Matute, D.R., and Cooper, B.S. (2021). Comparative studies on speciation: 30 years since Coyne and Orr. Evolution *75*, 764–778.

Morel, P., Heijmans, K., Ament, K., Chopy, M., Trehin, C., Chambrier, P., Rodrigues Bento, S., Bimbo, A., and Vandenbussche, M. (2018). The floral C-lineage genes trigger nectary development in *Petunia* and *Arabidopsis*. The Plant Cell *30*, 2020–2037.

Nakazato, T., Rieseberg, L.H., and Wood, T.E. (2013). The genetic basis of speciation in the *Giliopsis* lineage of *Ipomopsis* (Polemoniaceae). Heredity *111*, 227–237.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. Plant Cell *2*, 279–289.

Nosil, P., Crespi, B.J., and Sandoval, C.P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature *417*, 440–443.

Nouri, E., Surve, R., Bapaume, L., Stumpe, M., Chen, M., Zhang, Y., Ruyter-Spira, C., Bouwmeester, H., Glauser, G., Bruisson, S., et al. (2021). Phosphate suppression of arbuscular mycorrhizal symbiosis involves gibberellic acid signalling. Plant Cell Physiol *62*.

Ollerton, J., Winfree, R., and Tarrant, S. (2011). How many flowering plants are pollinated by animals? Oikos *120*, 321–326.

Orr, H.A. (2005). The genetic theory of adaptation: a brief history. Nature Reviews Genetics *6*, 119–127.

Quattrocchio, F., Wing, J., Woude, K. van der, Souer, E., Vetten, N. de, Mol, J., and Koes, R. (1999). Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color. The Plant Cell *11*, 1433–1444.

Ramsey, J., Bradshaw, H.D., and Schemske, D.W. (2003). Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (phrymaceae). Evolution *57*, 1520–1534.

Reck-Kortmann, M., Silva-Arias, G.A., Segatto, A.L.A., Mäder, G., Bonatto, S.L., and de Freitas, L.B. (2014). Multilocus phylogeny reconstruction: new insights into the evolutionary history of the genus *Petunia*. Molecular Phylogenetics and Evolution *81*, 19–28.

Reck-Kortmann, M., Mäder, G., Rodrigues, L.A., Ruas, C.F., and Freitas, L.B. (2017). AFLP markers contribute to species delimitation and evolutionary understanding of the recent genus *Petunia* (Solanaceae). Bot J Linn Soc *183*, 413–428.

Rieseberg, L.H., Whitton, J., and Gardner, K. (1999). Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics *152*, 713–727.

Rodrigues, D.M., Turchetto, C., Callegari-Jacques, S.M., and Freitas, L.B. (2018a). Can the reproductive system of a rare and narrowly endemic plant species explain its high genetic diversity? Acta Botanica Brasilica *32*, 180–187.

Rodrigues, D.M., Caballero-Villalobos, L., Turchetto, C., Assis Jacques, R., Kuhlemeier, C., and Freitas, L.B. (2018b). Do we truly understand pollination syndromes in *Petunia* as much as we suppose? AoB Plants *10*.

Särkinen, T., Bohs, L., Olmstead, R.G., and Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evolutionary Biology *13*, 214.

Schemske, D.W., and Bradshaw, H.D. (1999). Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). PNAS *96*, 11910–11915.

Schnitzler, C.K., Turchetto, C., Teixeira, M.C., and Freitas, L.B. (2020). What could be the fate of secondary contact zones between closely related plant species? Genet. Mol. Biol. *43*, e20190271.

Scopece, G., Musacchio, A., Widmer, A., and Cozzolino, S. (2007). Patterns of reproductive isolation in mediterranean deceptive orchids. Evolution *61*, 2623–2642.

Segatto, A.L.A., Cazé, A.L.R., Turchetto, C., Klahre, U., Kuhlemeier, C., Bonatto, S.L., and Freitas, L.B. (2014). Nuclear and plastid markers reveal the persistence of genetic identity: A new perspective on the evolutionary history of *Petunia exserta*. Molecular Phylogenetics and Evolution *70*, 504–512.

Shang, Y., Venail, J., Mackay, S., Bailey, P.C., Schwinn, K.E., Jameson, P.E., Martin, C.R., and Davies, K.M. (2011). The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. New Phytologist *189*, 602–615.

Sheehan, H., Moser, M., Klahre, U., Esfeld, K., Dell'Olivo, A., Mandel, T., Metzger, S., Vandenbussche, M., Freitas, L., and Kuhlemeier, C. (2016). *MYB-FL* controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nature Genetics *48*, 159–166.

Stehmann, J., and Semir, J. (2005). New species of *Calibrachoa* and *Petunia* (Solanaceae) from subtropical South America. Monographs in Systematic Botany from the Missouri Botanical Garden *104*, 341–348.

Stehmann, J.R., Lorenz-Lemke, A.P., Freitas, L.B., and Semir, J. (2009). The genus *Petunia*. In Petunia, (New York, NY: Springer New York), pp. 1–28.

Sucena, É., and Stern, D.L. (2000). Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of ovo/shaven-baby. PNAS *97*, 4530–4534.

Teixeira, M.C., Turchetto, C., Hartke, S., Schnitzler, C.K., and Freitas, L.B. (2019). Morphological and genetic perspectives of hybridization in two contact zones of closely related species of *Petunia* (Solanaceae) in southern Brazil. Acta Botanica Brasilica *33*, 734–740.

Teixeira, M.C., Turchetto, C., Maestri, R., and Freitas, L.B. (2020). Morphological characterization of sympatric and allopatric populations of *Petunia axillaris* and *P. exserta* (Solanaceae). Botanical Journal of the Linnean Society *192*, 550–567.

Turchetto, C., Segatto, A.L.A., Beduschi, J., Bonatto, S.L., and Freitas, L.B. (2015a). Genetic differentiation and hybrid identification using microsatellite markers in closely related wild species. AoB PLANTS *7*, plv084.

Turchetto, C., Lima, J.S., Rodrigues, D.M., Bonatto, S.L., and Freitas, L.B. (2015b). Pollen dispersal and breeding structure in a hawkmoth-pollinated Pampa grasslands species *Petunia axillaris* (Solanaceae). Ann Bot *115*, 939–948.

Vandenbussche, M., Chambrier, P., Rodrigues Bento, S., and Morel, P. (2016). Petunia, your next supermodel? Front Plant Sci *7*.

Watanabe, H., Ando, T., Tsukamoto, T., Hashimoto, G., and Marchesi, E. (2001). Cross-compatibility of *Petunia exserta* with other *Petunia* taxa. Journal of the Japanese Society for Horticultural Science *70*, 33–40.

Whittall, J.B., and Hodges, S.A. (2007). Pollinator shifts drive increasingly long nectar spurs in columbine flowers. Nature *447*, 706–709.

Widmer, A., Lexer, C., and Cozzolino, S. (2009). Evolution of reproductive isolation in plants. Heredity *102*, 31–38.

Xu, S., Schlüter, P.M., Grossniklaus, U., and Schiestl, F.P. (2012). The genetic basis of pollinator adaptation in a sexually deceptive orchid. PLoS Genet *8*, e1002889.

Yarahmadov, T., Robinson, S., Hanemian, M., Pulver, V., and Kuhlemeier, C. (2020). Identification of transcription factors controlling floral morphology in wild *Petunia* species with contrasting pollination syndromes. The Plant Journal *104*, 289–301.

Yuan, Y.-W., Sagawa, J.M., Young, R.C., Christensen, B.J., and Bradshaw, H.D. (2013). Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics *194*, 255–263.

Zhang, W., Reeves, G.R., and Tautz, D. (2021). Testing Implications of the Omnigenic Model for the Genetic Analysis of Loci Identified through Genome-wide Association. Current Biology *0*.

Chapter 2

Genetic architecture of colour and morphology in a pollinator shift between *Petunia axillaris* and *P. exserta*

Marta Binaghi[1], Korinna Esfeld[2], Therese Mandel[1], Loreta B. Freitas[3], Cris Kuhlemeier[1]

1. Institute of Plant Sciences, University of Bern, Bern 3013, Switzerland
2. Institute of Applied Physics, University of Bern, Bern 3012, Switzerland
3. Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Keywords: Petunia, pollination syndromes, evolution, secondary contact, morphology, GWAS.

Author contributions

C.K. and M.B. conceived the study, K.E. and T.M. collected phenotypic data and extracted DNA, L.B.F. provided the seeds, grew the plants, and reviewed the manuscript, M.B. performed the analyses and wrote the manuscript under the supervision of C.K.

Abstract

The definition of biological species is based on the presence of reproductive isolation between groups of organisms. Reproductive isolation between species can have different genetic bases and include several barriers. In *Petunia*, pollinators have an important role in the isolation of species that are not separated by geography, as in contact zones. The genetic basis of pollinator-specific traits can thus inform us on the molecular processes involved in speciation. We study individuals from two contact zones where the moth-pollinated *Petunia axillaris* and the hummingbird-pollinated *P. exserta* hybridise, in order to characterise the genetic architecture of the pollinator-specific traits, and to investigate the selection acting in these populations. We analyse a set of genomic variants in combination with phenotype data to estimate population structure, genetic architecture of the pollination syndrome phenotypes, and loci associated to these traits. We then identify regions under positive selection at the contact zones. We show that traits such as colour variation can have a complex genetic basis, while morphological traits such as reproductive organ lengths can be controlled by few genes of large and moderate effect. The selection scan reveals that the speciation gene *MYB–FL*, responsible for UV absorbance of the petals, is absent from the region under selection at the contact zones, and suggests that other phenotypic traits or genetic incompatibilities may be the target of selection in these populations. We thus conclude that morphological traits typically considered polygenic can be controlled by few genes and could be the target of selection during primary or secondary divergence as much as simpler traits associated to pollinator preferences.

Introduction

Under Mayr's Biological Species Concept, speciation is the evolution of reproductive isolation between diverging groups of individuals (Gao and Rieseberg, 2020). How barriers to reproduction arise and are maintained is therefore a main question in evolutionary genetics. In plants, reproductive isolation can be due to one or more interacting barriers limiting gene flow, for example geography, flowering time, or hybrid fertility (Baack et al. 2015). Among these barriers, the pollinator specificity of flowers towards one or a group of animals can strongly affect the total isolation between species (for example Carrió and Güemes, 2014; Ramsey et al., 2003). In these cases, floral traits that attract and reward the pollinator and ensure pollination efficiency are fundamental to the isolation of species (Bradshaw and Schemske, 2003; Bradshaw et al., 1995; Shang et al., 2011); such sets of traits specific to a pollinator are termed "pollination syndromes" (Fenster et al., 2004).

The genetic basis of pollination syndrome traits is studied to understand how reproductive isolation barriers can be established and maintained during speciation. Studies on the topic have found that different genetic architectures can be the basis of such traits (Bradshaw and Schemske, 2003; Byers et al., 2014; Fishman et al., 2002; Nakazato et al., 2013; Xu et al., 2012; Yuan et al., 2013). For the purpose of this manuscript we classify genetic architectures as: 1) monogenic, when one gene of major effect is involved and explains most of the phenotypic variation; 2) oligogenic, when less than 50 genes of major effect are involved and explain most of the phenotypic variation; 3) polygenic, when more than 50 genes are involved and their effect does not explain the majority of the phenotypic variation. Changes in floral pigmentation are usually considered as genetically simple, and in *Antirrhinum* venation patterning and in *Petunia axillaris* and *Petunia inflata* visible and UV colours have been shown to be underlain by changes in one gene (Esfeld et al., 2018; Hoballah et al., 2007; Quattrocchio et al., 1999; Shang et al., 2011; Sheehan et al., 2016). In other cases though, pigmentation showed an oligogenic and polygenic base such as in *Aquilegia* and in *P. exserta* visible colour

(Berardi et al., 2021; Edwards et al., 2021). Morphological traits on the other hand are generally considered to be polygenic due to their complex developmental processes. Several studies associated such traits to many loci that have a small effect on the phenotype (Fishman et al., 2002; Nakazato et al., 2013), but some examples of a monogenic basis with genes of large effect have been described as well (Chen et al., 2007; Edwards et al., 2021). It has to be noted that technical limitations restrain the ability to identify loci of small effect; consequently genes described in the literature are most often of large effect. The classical dichotomy in the genetic basis of quantitative and qualitative phenotypic traits is therefore not an absolute rule. The higher resolution provided by Next Generation Sequencing techniques nowadays allows the mapping of polygenic traits when the study is designed adequately, and empirical studies can thus generate a comprehensive understanding of the genetic basis of pollination syndromes and hence of reproductive isolation.

Petunia axillaris and *Petunia exserta* are closely related species that diverged 310 thousand years ago (Caballero-Villalobos et al., 2021). Their flowers are similar, but differ in traits related to pollination syndromes. *P. axillaris* is pollinated by hawkmoths and displays white, UV-absorbent flowers with partially fused petals forming a tube and a limb (Figure 2.1, Dell'Olivo and Kuhlemeier, 2013; Stehmann et al., 2009). The reproductive organs are as long as the tube and do not extend outside of it (Hermann et al., 2013). *P. exserta* instead is hummingbird-pollinated and displays red, UV-reflective flowers, whose reproductive organs extend outside of the petal tube, a typical trait in flowers pollinated by birds (Figure 2.1 B, Lorenz-Lemke et al., 2006). The loss of UV absorption in *P. exserta* is due to a frameshift mutation in the *MYB-FL* gene. The same gene was responsible for the gain of UV absorption in the shift from bee-pollinated *P. inflata* to moth-pollinated *P. axillaris* by means of the upregulation of its expression (Sheehan et al., 2016). The essentially monogenic architecture of this trait is opposed to the more complex change in visible colour from the white to the red flowers. Berardi and colleagues (2021) identified several genes involved in the gain of red colour in *P. exserta*, suggesting a more complex genetic basis underlying this trait. The

morphological difference in pistil length between the two species has instead been associated to a major QTL in chromosome 2 flanked by two other QTLs in chromosomes 5 and 7 (Hermann et al., 2015). A more recent study that focused on transcription factors identified two *EOB* (emission of benzenoids) genes that affect style length when silenced (Yarahmadov et al., 2020). These results point towards a more complex basis involving several genes in control of the pistil length difference between *P. axillaris* and *P. exserta* plants.

The genus *Petunia* is endemic to South America, where *P. axillaris* occurs in numerous and widespread populations, while *P. exserta* is only found in smaller groups in a few localities (Stehmann et al., 2009). *P. axillaris* is mostly self-incompatible and grows in open habitats, while *P. exserta* occupies small caves with shallow soil and limited sunlight, and the populations found in the wild have high levels of inbreeding (Segatto et al., 2014; Turchetto et al., 2015a, 2021). The two species are interfertile when artificially crossed, and hybrid individuals between the species are found in two contact zones (Caballero-Villalobos et al., 2021; Lorenz-Lemke et al., 2006; Schnitzler et al., 2020; Teixeira et al., 2019; Turchetto et al., 2019a; Watanabe et al., 2001). At these contact zones, small groups of *P. axillaris* and *P. exserta* plants grow as close as a few metres apart, and a few hybrid individuals display intermediate phenotypes for colour and morphology (Lorenz-Lemke et al., 2006; Schnitzler et al., 2020; Teixeira et al., 2020; Turchetto et al., 2019a). A recent study estimated that secondary contact between the species happened 920 years ago (99 % CI: 250 - 81,100 years ago) and is currently ongoing, and identified a higher migration rate from *P. axillaris* towards *P. exserta* (Caballero-Villalobos et al., 2021). In these contact zones hummingbirds occasionally visited *P. axillaris* plants, and may be responsible for gene flow across the species (Lorenz-Lemke et al., 2006). Even though hybridisation is ongoing, individuals displaying the typical *P. axillaris* and *P. exserta* phenotypes can still be found in the contact zones (Teixeira et al., 2019, 2020); this suggests that gene flow is limited and at least partial isolation between the species still exists (Turchetto et al., 2015a, 2021).

The complete evolution of a different pollination syndrome involves changes in several genes (Edwards et al., 2021). When two species with different syndromes come into secondary contact their genetic components can meet and recombine. If no barrier to this process exists, the phenotypic differences are lost or homogenised through the population. In the case of *P. axillaris* and *P. exserta* contact zones, parental-looking plants are still found, and display the canonical *P. axillaris* and *P. exserta* pollination syndrome traits. We therefore ask what the genetic architecture of the pollination syndrome trait differences between *P. axillaris* and *P. exserta* is, and whether the differences are controlled by few or many genetic changes. We then ask whether the loci responsible for these differences are still under selection in secondary contact. Pollinators are the main reproductive isolation barrier in these species when geography is excluded. The animal's behaviour and ability to transfer pollen effectively may therefore still apply selection pressure towards optimal floral traits in these populations. We therefore expect that the main pollination syndrome traits (UV and visible colour, and reproductive organ morphology) in these contact zones are under positive selection due to the pollinators' preferences and efficiencies.

Our results surprisingly show that the two discrete traits (UV and visible colour) are controlled by opposing genetic architectures (oligogenic and polygenic, respectively) and that neither of them is clearly under selection. The two morphological traits instead are controlled by an oligogenic architecture and are likely to be under directional selection in the secondary contact zones.

Results

The analysis is based on two contact zones between the species *P. axillaris* and *P. exserta* in the south of Brazil (Lorenz-Lemke et al., 2006; Schnitzler et al., 2020). We collected seeds from two locations, from all the 30 mother plants that were carrying seeds at the moment of our visit. We grew the progeny from seeds in a greenhouse to minimise environmental effects. A set of 70 plants was selected in order to maximise phenotypic variation (detailed description in Materials and methods). The final set of 70 plants includes 60 from seeds collected in the first contact zone and 10 from contact zone 2. Genomic DNA of the individual plants was sequenced. The reads were mapped to the reference genome of *P. axillaris* and had an average coverage of 5× and a minimum coverage of one read on 79 % of the genome length. The variant calling produced a set of 4,167,183 biallelic positions with call rate higher or equal to 90 % and minor allele frequency (MAF) higher or equal to 0.05. The uncertainty in genotype calls caused by the low coverage sequencing was accounted for by using genotype likelihoods where the software allowed it.

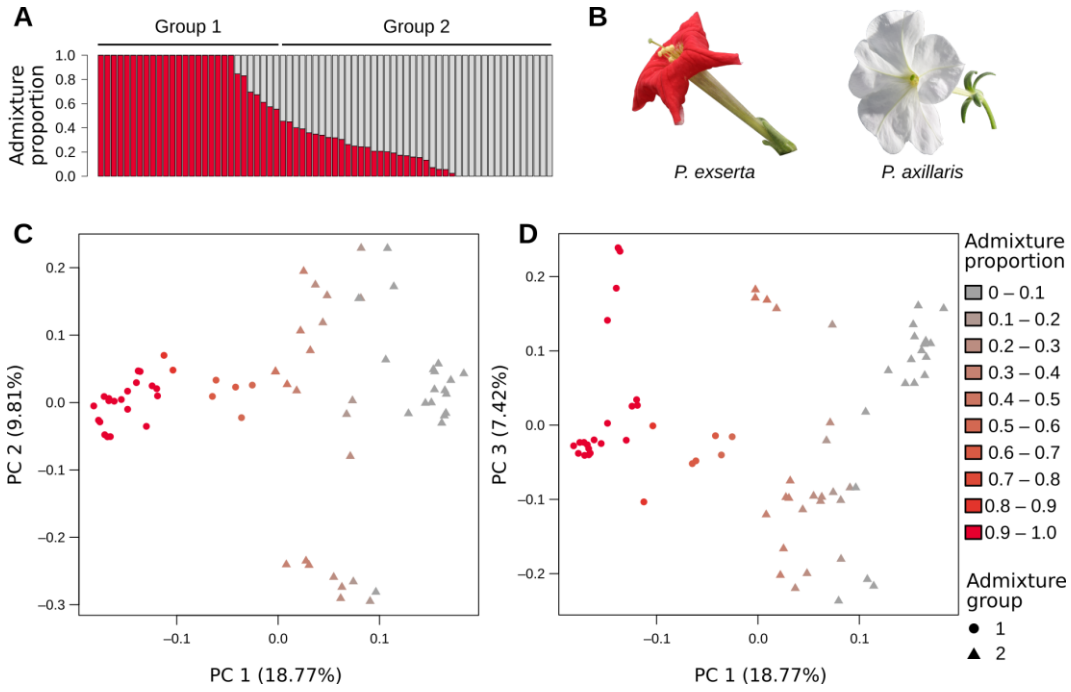


Figure 2.1. Population genetics of the wild *P. axillaris* × *P. exserta* hybrids. A) Genomic admixture proportion of the analysed individuals for $K = 2$. B) Photos of *P. exserta* (left) and *P. axillaris* (right) flowers. C and D) Genomic PCA of the analysed individuals. In panel C the first and second principal components are compared, in panel D the first and third principal components. The colour represents the individual admixture proportion assigned for $K = 2$, the symbol represents the admixture groups as defined on a threshold of 0.50 (group 1 ≥ 0.50, group 2 < 0.50).

Population structure analyses reveal admixed individuals at different generations

The plants included in our experiments can have different genetic histories: they can be pure species, they can be interspecific hybrids, or backcrosses or higher-order hybrids. It is therefore essential to describe the genomic structure of the population. For this purpose we considered all the individuals from the two different contact zones together, due to the low number of individuals present in one of these zones. We performed an admixture analysis for K (number of putative ancestral populations) between 1 and 8. We hypothesised that the most likely K would be 2, representing the two species. The admixture analysis was performed on the genotype likelihoods of the sites with call rate of 90 % or higher, and MAF ≥ 0.05. Evanno's method identified $K = 2$ as the most likely number of clusters (Figure 2.1 A and Supplementary file S1, Evanno et al., 2005). Of the 70 individ-

uals, 31 have admixture proportions lower or equal to 0.25 and 23 have admixture proportions higher or equal to 0.75. We use thresholds to define our pure species based on previous simulations and adjusted for a sufficient number of individuals in the two classes (Schnitzler et al., 2020; Segatto et al., 2014). The remaining 16 individuals are admixed, and represent plants derived from a cross between the two species. In total, 28 individuals belong to admixture group 1 (proportion higher or equal to 0.50) and 42 to group 2 (admixture lower than 0.50). Group two is therefore more numerous in our populations. These results confirm the presence of pure species and admixed individuals in these zones, corroborating the latest findings (Caballero-Villalobos et al., 2021; Schnitzler et al., 2020).

The results of the admixture analysis are confirmed by a principal component analysis (PCA) of the genomic data (Figure 2.1 C, D). The first principal component (PC) accounts for 18.77 % of the genetic variation and correlates with the admixture proportion for $K = 2$ (Spearman's rank correlation rho 0.98, P value lower than 2.2×10^{-16}). The second principal component accounts for 9.81 % of the genetic variation and is mostly represented in individuals of admixture group 2 (Figure 2.1 C), suggesting that this group is more diverse than group 1, and has a "private" axis of genetic variation. The third principal component accounts for 7.42 % of the variation and is represented in both admixture groups (Figure 2.1 D). This axis of genetic variation is partly associated with the different location of origin of the individuals (Supplementary file S2).

Table 2.1 Phenotype summary values and normality

Trait	Whole population		Admixture ≤ 0.25		Admixture ≥ 0.75		Normality *W* (*P*)
	Median	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	
Flavonol (AU)	56.66	49.53 (27.70)	61.58	54.48 (18.14)	21.98	42.69 (35.06)	0.92 (0.0003)
Antho. (AU)	2.33	11.48 (16.32)	0.47	0.70 (0.79)	33.91	31.19 (14.83)	0.72 (3.6×10^{-10})
Tube l. (cm)	4.66	4.69 (0.42)	4.78	4.80 (0.38)	4.62	4.53 (0.50)	0.98 (0.18)
Pistil l. (cm)	5.03	5.13 (0.50)	4.96	5.04 (0.45)	5.36	5.28 (0.61)	0.98 (0.23)

AU, absorbance unit
Antho, anthocyanin content
Tube l, tube length
Pistil l, pistil length

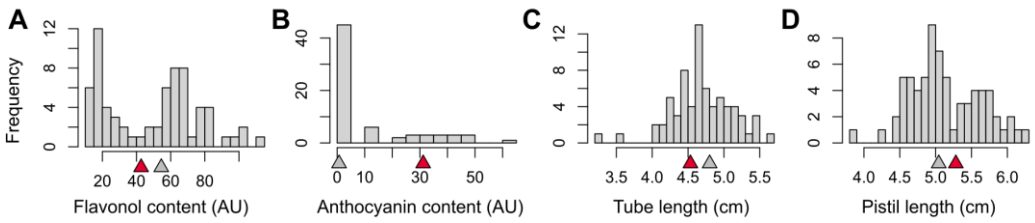


Figure 2.2 Distributions of the phenotypic values of the individuals included in the genetic analysis. The grey and red triangles represent the mean of the individuals with admixture proportion ≤ 0.25 and ≥ 0.75 respectively.

Phenotype distributions reflect the discrete and continuous architectures of the traits

We measured the phenotypic traits responsible for visible and UV colour (anthocyanin and flavonol content) and the traits responsible for the position of the stigma relative to the petal surface (tube and pistil length).

The presence of flavonols determines the UV absorption of petals and provides guidance for hawkmoths when they approach flowers (Sheehan et al., 2016; White et al., 1994). The flavonol content is generally high in *P. axillaris* and low in *P. exserta*, and artificial interspecific crosses (such as F_1 and F_2) generally show values in the same range of one of the parents. Intermediate values for flavonol

content are therefore less common (Hermann et al., 2013). The flavonol content in our samples is not normally distributed and shows a bimodal trend, consistent with the monogenic architecture of the trait previously described (Figure 2.2 A, Table 2.1. Hermann et al., 2013; Sheehan et al., 2016). We compared the means of the individuals representative of the pure species (admixture proportion lower or equal to 0.25 and higher or equal to 0.75). The flavonol content is not different between the two groups (Wilcoxon rank sum test $W = 261$, $P = 0.1524$), suggesting that this difference between species is not strongly maintained in the two admixture groups at the contact zones. The contact zones display bimodal variability in flavonol content consistent with previous results and with the architecture of the trait. Furthermore the trait variability does not seem strongly associated with the genetic component of the admixture groups.

Anthocyanin content shows a bimodal distribution with a mass point at 0, suggesting that the trait could be controlled by one or few genes of major effect (Figure 2.2 B). The difference between the groups representing the species is statistically significant (Wilcoxon rank sum test $W = 681$, $P = 8.648 \times 10^{-15}$) and reflects the anthocyanin content previously reported in the species: *P. axillaris* has white flowers and does not accumulate these pigments, while the red colour of *P. exserta* is composed of anthocyanins (Berardi et al., 2021; Hermann et al., 2013). In the contact zones analysed, the difference in anthocyanin content between the admixture groups suggests that the trait is associated to the distinctive genetic components of each species. This observation is consistent with previous results that showed that visible colour at the contact zones is a good indicator of the genetic origin of each individual (Teixeira et al., 2019). Visible colour is therefore a strongly bimodal trait, which suggests a simple genetic architecture. Its association with the admixture groups suggest that the colour differences are caused by loci strongly associated to the species' genetic components. This observation does not exclude the possibility that the trait is controlled by few loci distributed along the genome.

Tube length is normally distributed, and the means of the admixture groups show a significant difference (Welch two sample t test $t = -2.1457$, $P = 0.03815$. Figure 2.2 C, Table 2.1). The distribution of the trait is consistent with a polygenic architecture of the trait, and the difference between admixture groups suggests that the trait variation could be maintained in association with the genetic components of the species.

Pistil length is normally distributed similarly to tube length (Table 2.1, Figure 2.2 D). The means of the two admixture groups are not different (Welch two sample t test $t = 1.5829$, $P = 0.1215$), although pistil length was previously described as longer in *P. exserta* accessions (Hermann et al., 2015). The lack of significant difference between the admixture groups suggests that the trait is not maintained in association with the species genetic components in these contact zones, possibly due to the introgression process eroding the phenotypic diversity. This result is supported by the findings of Turchetto and colleagues (Turchetto et al., 2019a), who found that *P. axillaris*-looking individuals growing in the typical *P. axillaris* habitat could have experienced introgression from *P. exserta*.

The distributions of the four traits are consistent with their classical categorisation as discrete (for anthocyanin and flavonol content, responsible for the colour) and continuous traits (for pistil and tube length). The pairwise correlations between phenotypic traits reveals only one statistically significant positive correlation between pistil and tube length (Supplementary file S3), probably motivated by their shared developmental processes. Our results indicate that visible colour and tube length are more strongly associated with a pure species genetic origin than UV absorbance and pistil length.

The pollination syndrome traits of Petunia *have different genetic architectures*

The genetic architecture of the phenotypic traits was estimated with the Bayesian Sparse Linear Mixed Model (BSLMM) provided in GEMMA . This association model assumes that all variants have at least a small effect, and that a certain pro-

portion of the variants has an additional effect (called “sparse” or “major” effect); both the effect parameters are derived from a normal distribution, and can take the value of 0 (Zhou et al., 2013). Thus, the application of this model allows for the correct modelling of traits controlled by few variants of major effect and additional small effect variants, as well as traits controlled only by numerous small effect variants.

Table 2.2 Predicted genetic architecture of the phenotypic traits

Trait	PVE[a] Mean (SD)	PGE[b] Mean (SD)	N gamma[c] Mean (SD)
Flavonol content	0.99 (0.02)	0.83 (0.09)	5.40 (4.03)
Anthocyanin content	0.99 (0.01)	0.35 (0.34)	71.10 (78.82)
Tube length	0.96 (0.07)	0.74 (0.15)	7.73 (8.81)
Pistil length	0.93 (0.09)	0.66 (0.24)	33.25 (56.80)

[a] Proportion of variance explained by genetic data
[b] Proportion of variance explained by sparse effect variants
[c] Number of variants with sparse effect

The four traits have very high variance explained by the genetic data (PGE, above 0.90), indicating that the trait is highly heritable and is not influenced by the environment in controlled conditions (Table 2.2). The proportion of variance explained by sparse effect variants (PVE) instead is diverse across the traits (Table 2.2). This proportion represents the variance in phenotype that can be attributed to variants with a “major” or “sparse” effect, and implicates that the remaining proportion (1 − PGE) is explained by variants with very small effect. Flavonol content has the highest PGE (0.83, Table 2.2), and the number of variants that are predicted to be responsible for that proportion (N gamma) is 5.40 on average. This result is consistent with the monogenic control of the trait already described in the species (Sheehan et al., 2016). Tube length has the second highest PGE (0.74) and an average N gamma of 7.73. This result suggests that tube length is mostly controlled by a few variants with major effects, although 0.26 of the proportion in phenotypic variance is explained by small effect mutations. Pistil length has a PGE of 0.66

and average N gamma of 33.25, indicating a rather intermediate architecture: more than half the phenotypic variance is controlled by sparse effect variants, but the number of these is higher than for the previous traits. Anthocyanin content has the lowest PGE with 0.35 and an N gamma of 71.10 (Table 2.2). These results point towards a polygenic architecture for this trait, with as many as 71 positions responsible for only 0.35 of the phenotypic variation. This result seems to contradict the findings of Berardi and colleagues (2021), who suggested an oligogenic control of the trait. Nonetheless, the predicted architecture could be explained by the strong association of this phenotype with the major axis of genetic variation in the populations (Supplementary file S3). This axis of variation is associated with the two species and presents the highest divergence; hence a phenotypic trait that variates along the same axis will be most likely associated to the species' genotypic divergence.

The traits considered therefore have different genetic architectures; these do not necessarily follow the classical understanding that quantitative traits such as pistil and tube length are controlled only by numerous mutations of small effect, while qualitative traits such as colour are controlled by mutations of large effect. Instead, visible colour could be controlled by several loci, suggesting that the evolution of this trait has followed several small to medium effect changes, and tube length variation could be controlled by few loci, despite its classical polygenic classification.

Association analysis identifies sites under selection associated to pistil and tube length variation

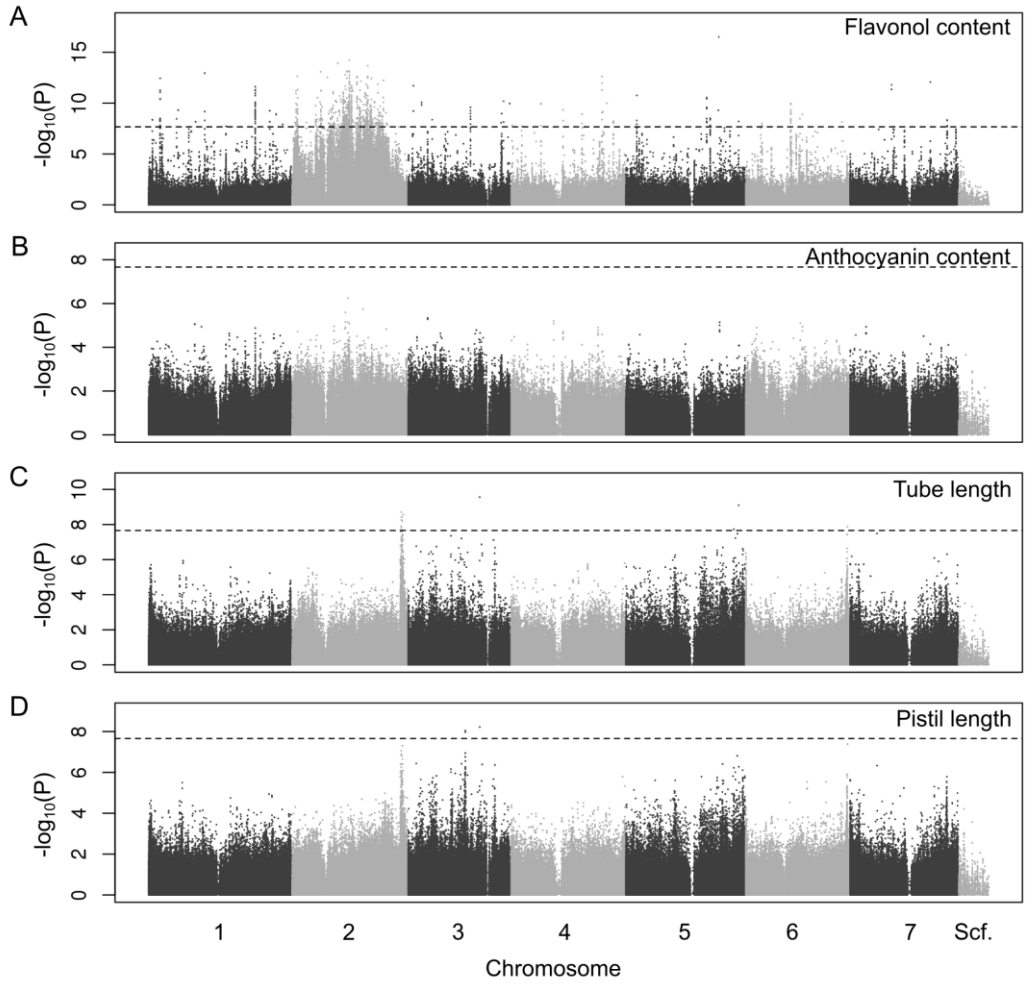


Figure 2.3 Genome-wide association study shows different regions associated to the pollination syndrome traits. Manhattan plots of the linear mixed model GWAS analysis. Dots represent the $-\log_{10}$ transformed *P* values of each variant. The horizontal dashed lines represent the Bonferroni-corrected threshold for $P = 0.05$. Scf indicates scaffolds not anchored to the chromosomes in the genome assembly.

We performed genome-wide association analyses for each phenotypic trait. We used a linear mixed model to account for population structure, but we note that the small sample size of our study limits the power of the association analysis; in particular, the statistical significance of loci of small to medium effect is hampered by the strict threshold imposed by the high number of variants considered. The

association analyses reveal different patterns across the phenotypic traits (Figure 2.3). Flavonol content has several significant sites, most of them situated on chromosome 2 (Figure 2.3 A). This result may seem to contradict the genetic architecture predicted and is discussed in a later section (Table 2.2). Anthocyanin content does not reveal any significant peak, suggesting that our study is underpowered for the identification of variants associated to this phenotype (Figure 2.3 B). This trait has a predicted polygenic basis, thus the effect of each variant could be small and require high statistical power to be detected (Table 2.2). Tube length has a significant peak at the end of chromosome 2, and a few sparse significant variants on chromosome 3 and 5 (Figure 2.3 C). Chromosome 3 and 5 association signals are distributed along the chromosomes and are mostly not significant, although it has to be noted that they are generally higher than the signal on the other chromosomes. Pistil length has only a few significant sites on chromosome 3; this chromosome shows a similar pattern to the association for tube length (Figure 2.3 D). Pistil length seems also to be associated to sites at the end of chromosome 2, in the same region identified for tube length. In this region the sites do not pass the Bonferroni-corrected threshold but form a clear peak that stands out from the rest of the chromosome. The different pollination syndrome traits analysed are therefore associated to restricted regions of the genome, except for anthocyanin content for which we were unable to identify associated loci.

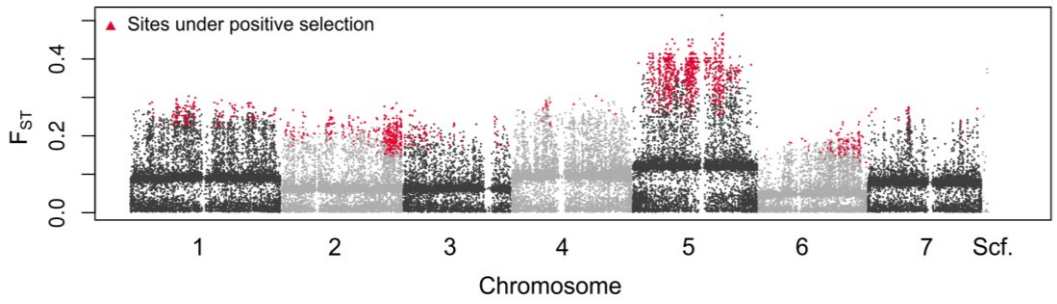


Figure 2.4 Selection scan identifies several regions of the genome under positive selection. Selection scan between admixture groups (<0.50 proportion, ≥ 0.50 proportion). Each dot represents a 10 kb window, red triangles are statistically significant for being under positive selection (F_{ST} > 0.15 and FDR < 0.05). Scf indicates scaffolds not anchored to a chromosome in the genome assembly.

To identify if any region of the genome is under selection in the contact zones we performed a selection scan comparing the two admixture groups (proportion threshold 0.50). This method allows the identification of loci where the fixation index between two groups (F_{ST}) is significantly different from that expected under neutral theory (Foll and Gaggiotti, 2008). The method identifies sites under positive selection as those with a high F_{ST}, and provides a posterior probability for each site for being under selection. The selection scan identifies several sparse sites under positive selection and two bigger regions on chromosome 2 and 5 (Figure 2.4). Chromosome 5 has high F_{ST} along most of its length, and its signal is the highest across the genome. This result suggests that chromosome 5 is experiencing strong selection pressure at the contact zones. Chromosome 2 instead shows a region under selection on its end; this region includes the sites associated to tube and pistil length in the GWAS analyses (Figure 2.5).

We found significant association signals for flavonol content and tube length on two different regions of chromosome 2 (Figure 2.5). The signal for flavonol content is extended along the chromosome and includes the gene *MYB-FL*, known to control the difference in UV absorbance between *P. axillaris* and *P. exserta* (Sheehan et al., 2016). On the same chromosome but outside of the region associated to flavonol content, tube and pistil lengths show a localised signal suggesting that one or more genes in that region affect these two traits (Figure 2.5 B, C). The same region shows high F_{ST} values when individuals with extreme phenotypes are compared (groups defined by the first and last quartile of the distribution of the trait. Figure 2.5 E, F). The region at the end of chromosome 2 is under selection when individuals with admixture proportion below 0.50 are compared to those of proportion higher than 0.50. In addition the selection scan shows that most of chromosome 5 is under selection (Figure 2.4).

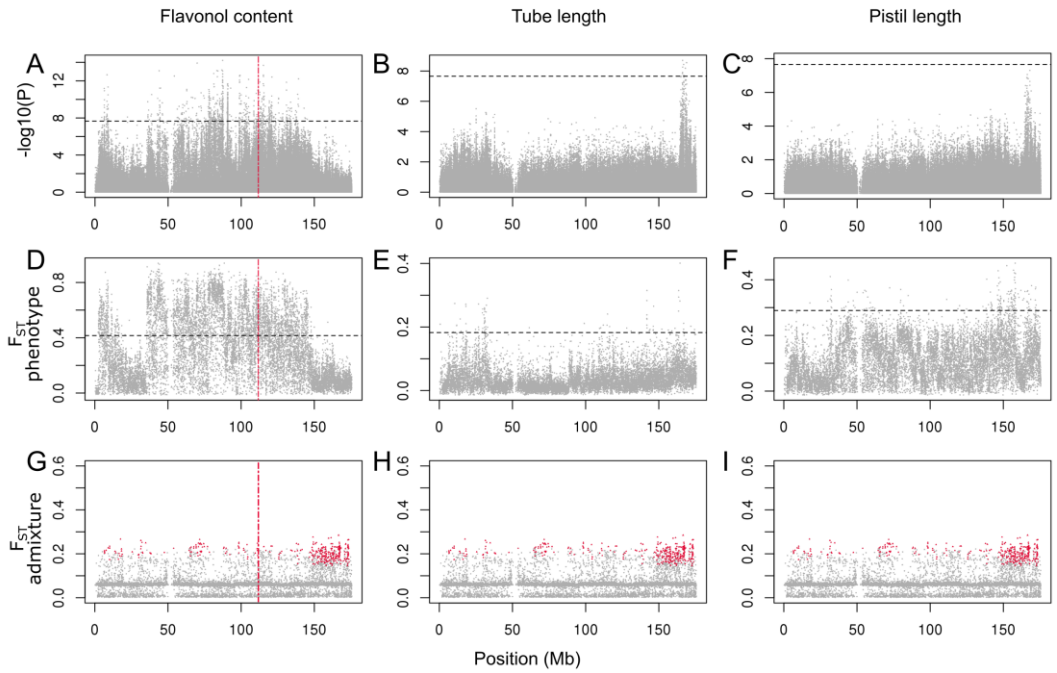


Figure 2.5 Morphological traits are associated to a region under positive selection on chromosome 2. A – C) Manhattan plots. Horizontal dashed lines represent the Bonferroni-corrected threshold for $P = 0.05$. D – E) F_{ST} between groups of individuals with different phenotypes (phenotype quantiles < 0.25 VS > 0.75). Each dot represents a 10 kb window. Horizontal dashed lines represent the 95th quantile. G – I) Selection scan between admixture groups (<0.50 proportion, > 0.50 proportion). Each point represents a 10 kb window. Red points are statistically significant for being under selection ($F_{ST} > 0.15$ and FDR < 0.05). The panels G to I are repeated to allow comparison with the panels above. A, D) flavonol content. B, E) tube length. C, F) pistil length. The red vertical lines represent the position of *MYB-FL*.

Discussion

P. axillaris and *P. exserta* produce viable offspring when artificially crossed, and hybrids between the two species are found in two contact zones in the South of Brazil (Caballero-Villalobos et al., 2021; Lorenz-Lemke et al., 2006; Watanabe et al., 2001). The reproductive isolation between the two species is therefore not complete; at the contact zones, hybridisation has been ongoing for an estimated 920 years, and individuals present different levels of genetic admixture (Caballero-Villalobos et al., 2021). Our analysis of the genetic components of individuals from two contact zones reveals one major axis of genetic variation likely associated to the ancestral species *P. axillaris* and *P. exserta*. Based on the phenotype of individuals with low admixture we suggest that admixture group 1 includes plants with a majority of *P. exserta* genetic ancestry and group 2 includes plants with a majority of *P. axillaris* ancestry (Figure 2.2, Table 2.1). The second axis of genetic variation is private to admixture group 2, therefore supporting the idea that group 2 is derived from *P. axillaris* (Figure 2.2 C). This species is geographically widespread and genetically more diverse than *P. exserta* that shows higher levels of inbreeding and a small effective population size (Schnitzler et al., 2020; Segatto et al., 2014; Turchetto et al., 2021).

P. axillaris and *P. exserta* are pollinated by hawkmoths and hummingbirds respectively, and the floral traits associated to the respective pollinators may constitute a reproductive barrier between the species (Dell'Olivo and Kuhlemeier, 2013; Dell'Olivo et al., 2011; Stehmann et al., 2009). We found that the pollination syndrome traits in these species have different genetic architectures. In our analysis, flavonol content is predicted to be oligogenic, with 83 % of the phenotype variation being controlled by 5 variants (Table 2.2). This prediction is consistent with the previously described major role of *MYB-FL* in the control of UV absorbance in these species (Sheehan et al., 2016). On the other hand, the GWAS result indicates that most of chromosome 2 is associated to the phenotype (Figure 2.3). This pattern can be explained by the reduced recombination previously observed in this

chromosome in *P. axillaris* × *P. exserta* crosses, and it is confirmed by an analysis of linkage disequilibrium performed on the individuals from the contact zones (Tenreira et al. in preparation, Hermann et al., 2013). The variation in UV colour at the contact zones is therefore most likely associated to the previously reported speciation gene *MYB-FL*.

The anthocyanin content is a major player in the definition of red colour in *P. exserta* (Berardi et al., 2021). We found that anthocyanin content was significantly different between lowly admixed individuals (admixture proportion ≤ 0.25 or ≥ 0.25), and its genetic basis was predicted to be polygenic (Table 2.2). The trait difference between the admixture groups suggest that colour is strongly associated with the species' genetic components, consistently with previous results by (Teixeira et al., 2019, 2020). The predicted polygenic basis of the trait is consistent with the GWAS result which displays no significant signal, possibly because of the small effect of the loci involved (Figure 2.3, Table 2.2). On the other hand, a recent study suggests that the shift to red colour in *P. exserta* is due to multiple and subtle genetic alterations, indicating an oligogenic architecture (Berardi et al., 2021). Our findings do not exclude this possibility because the association of the trait with the admixture groups and hence with the genetic structure of the population could conceal the associated loci in the GWAS. If the trait is strongly associated to the genetic diversity that species have accumulated during divergence, its causative loci should be associated to many more species-specific loci along the genome. The correction for population structure in the GWAS analysis would therefore mask the true causative loci.

Tube length difference between admixture groups is not strongly significant, and pistil length difference is not significant, suggesting that these traits are not strongly associated to the species' genetic differentiation. The genetic architecture of the traits suggests an oligogenic basis, with pistil length controlled by more variants of sparse and of small effect than tube length (Table 2.2). Accordingly with the smaller effect predicted for variants controlling pistil length, this trait has fewer loci passing the association test (Figure 2.3). We note that the region at the

end of chromosome 2 associated to tube length is also showing a signal for association in pistil length, and that chromosome 2 had already been associated to pistil and tube length in a previous QTL study (Hermann et al., 2015). The colocation of the tube and pistil length signals could be due to the same causative variants because of the shared underlying processes that control floral organ development. Yarahmadov and colleagues (2020) showed that the transcription factors EOBI and EOBII affect pistil length in both *P. axillaris* and *P. exserta*, and they showed that the MYB class is overrepresented among the transcription factors associated with pistil length variation. The involvement of a transcription factor with pleiotropic effects could explain the colocation of the signal on chromosome 2.

Our selection scan identified a wide region on chromosome 5 and a region at the end of chromosome 2 as being under positive selection (Figure 2.4). These signals could be due to the ancient divergence process that led to speciation or to a more recent action of selection at the secondary contact zones. In particular, the presence of incompatibilities between the species could show a similar signal in a young hybrid population (Lindtke and Buerkle, 2015). The comparison of selection scans obtained between sympatric and allopatric pairs of populations could better inform us on the role of selection pressure between these two species, indicating if selection acts in the divergence of species or only in their maintenance at the contact zones.

We hypothesise that pistil and tube length could be experiencing selection pressure at the contact zones, while UV absorbance is not. The region under selection on chromosome 5 does not co-localise with any of the phenotype-associated regions. On chromosome 2 instead the region under positive selection overlaps with the association signal for pistil and tube length (Figure 2.5). The importance of UV and the role of *MYB-FL* in speciation is not to be excluded though. This trait could have been important in the speciation of *P. exserta*, when a new pollination syndrome had to be established quickly. The morphological differences of the flower instead could have been developed later during the divergence of the two species. The position of the floral organs can affect the transfer of pollen to and

from the pollinator, and could therefore represent a barrier when the same pollinator visits both species (Wolf et al., 2001).

Future experiments will help to confirm the role of pistil and tube length in reinforcing the reproductive barriers between *P. axillaris* and *P. exserta*. The regions identified by the GWAS and selection scan should be investigated in detail to identify the genes underlying the phenotypic differences, and pollinator assays could be used to evaluate the effect size of such differences. Our work suggests that morphological traits typically considered polygenic can be controlled by few genes, and that morphology of flowers could constitute a trait that is selected during the species divergence or upon secondary contact between flowers pollinated by hummingbirds and hawkmoths. Understanding these patterns of selection among traits thus helps us to better understand how speciation occurs and is maintained as well as what patterns may or may not hold across wider ranges of organisms experiencing similar selective pressures.

Materials and methods

Plants

Seeds were collected in two contact zones of *P. axillaris* subsp. *axillaris* and *P. exserta* in the Guaritas region of Brazil, from all plants carrying seeds at the time of our visit (November 2011). These contact zones have been described in Caballero-Villalobos et al., 2021; Lorenz-Lemke et al., 2006; Schnitzler et al., 2020; Segatto et al., 2014; Teixeira et al., 2019, 2020; Turchetto et al., 2015b, 2019b. Contact zone 1 corresponds to coordinates 30°53′48″ S, 53°25′16″ W and contact zone 2 corresponds to 30°50′14″ S, 53°30′15″ W. The seeds were sown in pots and grown in a greenhouse, in order to limit environmental effects on the phenotypes.

Measurement of floral phenotypic traits and selection of plants

Between 3 and 8 flowers per plant were used for phenotype recording. Flowers were collected two days post anthesis and the corolla was cut open longitudinally. The flower was then pinned to a flat surface and photographed. Photos were processed in Fiji (ImageJ) to obtain pistil and tube length (Schindelin et al., 2012). Pistil length was measured from the base of the ovary to the top of the stigma. Tube length was obtained by summing the length of the petal tube sections D1 and D2. D1 limits are defined by the petal attachment point at the base of the ovary and the point where the filaments of the stamen detach from the corolla. D2 starts where the filaments detach and ends where the corolla bends and constitutes the limb of the petals. The average between flowers was taken to represent the individual plant. After measuring the morphological traits, a disc of 8 mm in diameter was sampled from the corolla limb of the flower. The measurement of the pigments was performed with a spectrophotometer after extraction, as described in Sheehan et al., 2016. Recorded values were averaged to obtain the plant measurement.

All following analyses were performed on a subset of plants selected among the greenhouse-grown individuals. In order to maximise the phenotypic diversity we

selected individuals displaying different combinations of traits, e.g. a plant with red flowers but a short pistil, or a plant with white and UV-reflective flowers. In order to maximise genetic diversity we selected plants from most mothers we had available (28 out of a total of 30). Per plant phenotype measurements are available in Supplementary file S4.

DNA extraction and sequencing

Leaf tissue was collected from each plant and DNA was extracted with a modified CTAB protocol (Murray and Thompson, 1980). The DNA was quantified with a fluorometer (Invitrogen Qubit™). Library preparation and sequencing were performed by the Next Generation Sequencing platform of the University of Bern in two batches, one in 2016 and one in 2018. In both batches, the DNA was amplified with illustra™ GenomiPhi™ V2 DNA Amplification Kit. Library preparation followed the TruSeq DNA PCR-free protocol. Sequencing was performed to obtain 150 bp long, paired-end reads, for an estimated coverage of 4-5 ×, on a genome size of 1.2 Gb (Bombarely et al., 2016). The 2016 batch was sequenced on two lanes of an Illumina HiSeq 3000. The 2018 batch was sequenced on two lanes (one chip) of the Illumina NovaSeq.

Read alignment, variant calling and filtering

Raw reads were quality controlled with FastQC version 0.11.7 (Andrew, 2010). Trimming and adapter removal was performed with Trimmomatic version 0.36 (Bolger et al., 2014). Reads were aligned to the reference sequence of *P. axillaris* version 4.0.2 (in preparation). Alignement was performed with BWA-MEM 0.7.17, and lanes were merged with samtools 1.10 (Danecek et al., 2021; Li and Durbin, 2009) at bam file stage. A coordinate file was produced to exclude from variant calling the repetitive regions of the genome, and the regions with a coverage higher than 100 reads in one or more samples. GATK 4.1.3.0 and tools included in it (McKenna et al., 2010) were used to mark duplicated reads and to perform variant calling. The obtained variants were hard filtered following the GATK best practices for organisms that lack panels of high-quality variants. Before applying the GATK suggested thresholds, the quality parameters were extracted for

SNPs and INDELs and their distribution was observed to confirm that filter thresholds were appropriate. After quality filtering, variants were then filtered to keep only positions with 90% or higher call rate, and minor allele frequency of 0.05 or higher.

Population genetics and GWAS analyses

Population genomics analyses (PCA and admixture) were performed on genotype likelihoods rather than on called genotypes, to account for the uncertainty introduced by the low coverage on the genotype calls. Genotype likelihoods were calculated from the PL field in the vcf file produced by GATK with a custom Python script available on the GitHub page associated to this manuscript. The genotype likelihoods were then used to perform an admixture analysis with NgsAdmix available in ANGSD version 0.933 (Korneliussen et al., 2014; Skotte et al., 2013). *K* from 1 to 8 were tested, and the analysis repeated 10 times. The run with the best likelihood was chosen for each *K*. The most likely *K* was calculated using Evanno's method (Evanno et al., 2005) implemented in an R script. PCA analysis was performed with pcangsd version 1.02 (Fumagalli et al., 2013). F_{ST} between groups of individuals with different phenotype values was calculated with ANGSD.

The GWAS and the prediction of the genomic architecture of the phenotypic traits was performed with GEMMA version 0.98.4 (Zhou and Stephens, 2012). The vcf file was converted with a modified version of a Perl script by Victor Soria-Carrasco available at https://github.com/visoca/popgenomworkshop-gwas_gemma. The phenotype files were formatted according to GEMMA guidelines in R.

The genetic architecture prediction was performed with the Bayesian sparse linear mixed model (BSLMM) provided in GEMMA, with options `-bslmm 1` (fits a linear BSLMM using MCMC), `-w 50000000` (discarded burn-in iterations), `-s 200000000` (saved sampling iterations) (Zhou et al., 2013). The mean and SD of each hyperparameter were calculated in R.

The GWAS was performed with the univariate linear mixed model, option `-lmm 4`, and the likelihood ratio test was considered for significance. *P* values were corrected for multiple testing by dividing the canonical 0.05 threshold by the number of variants tested. Manhattan plots were produced in R.

The selection scan was performed in R, with the function BayeScanR provided in the package PopGenome (Foll and Gaggiotti, 2008; Pfeifer et al., 2014). The function implements in R the classical method by Foll and Gaggiotti.

Resource-demanding computations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern. Statistical data analyses and plotting were performed in R (R Core Team, 2021; RStudio Team, 2015).

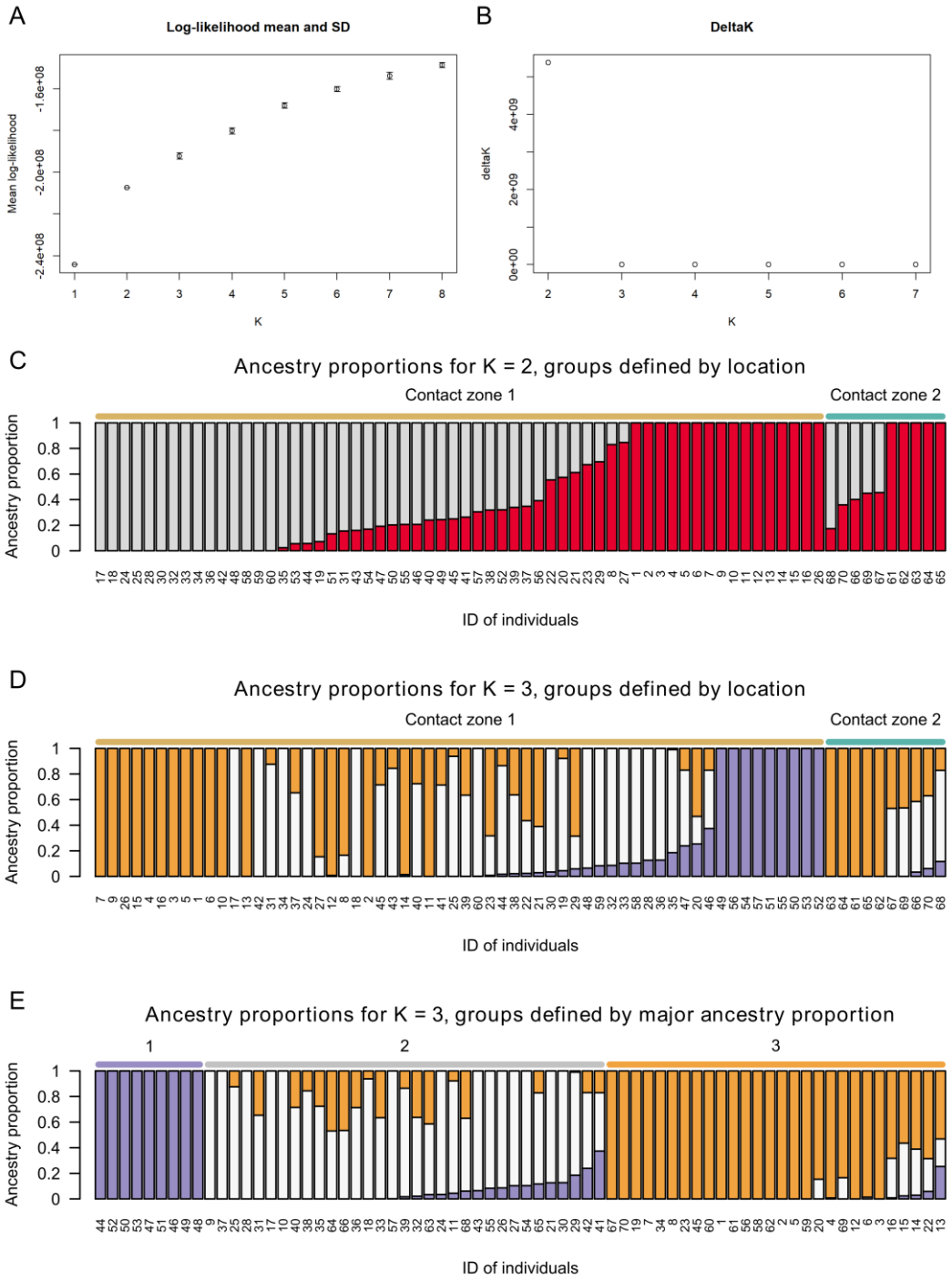
Data and code availability

Raw reads are available on NCBI SRA, under the BioProject accessions PRJNA522653 (2016 batch) and PRJNA706535 (2018 batch). Scripts and parameters used in the analyses are available on GitHub, at https://github.com/Kuhlemeier-lab/genetics_petunia_hybrids.

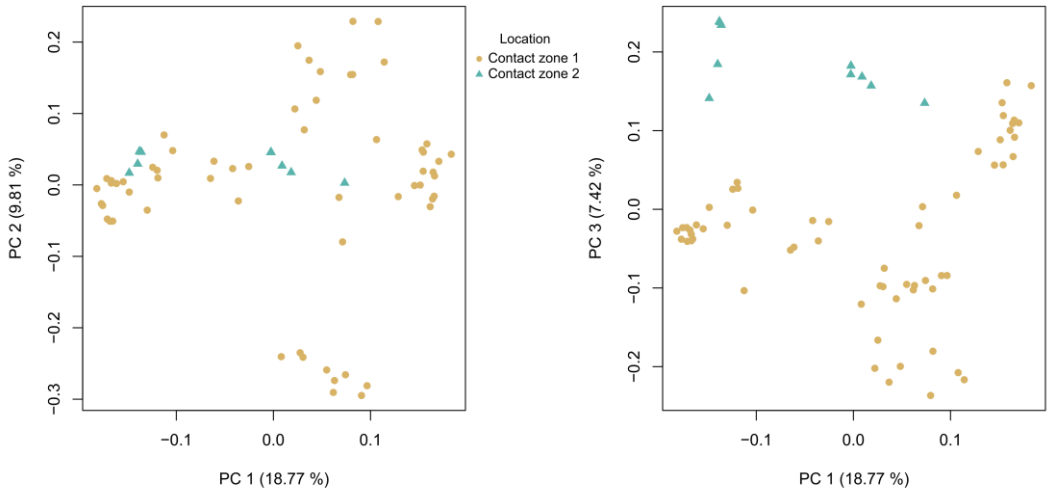
Acknowledgements

We thank Andrea Berardi, Mathieu Hanemian, Dorothea Lindtke, Martina Lüthi, Michel Moser, and Tracey Tenreira for technical support and helpful discussion. We thank Kimberly J. Gilbert for her precious advice on population genetics matters. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant agreement No. 741354 and the Swiss National Science Foundation (Grants 31003A_159493 and 31003A_1823403 to C.K.), and from The Brazilian National Council for Scientific and Technological Development (Grants to L.B.F.).

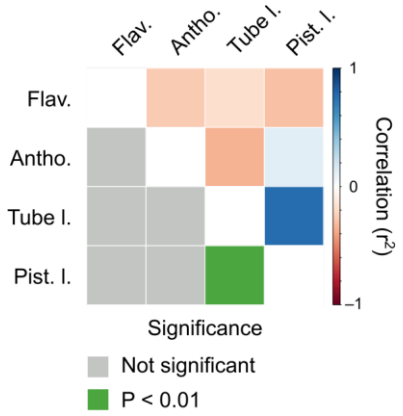
Supplementary data



Supplementary figure 1 Evanno's choice of *K* and admixture proportions for *K* 2 and 3. A) Log-likelihood mean and standard deviation of each *K* tested in NGSADMIX. B) Delta of the log-likelihoods. C) Admixture proportions for $K = 2$, with individuals sorted by contact zone of origin. D, E) Admixture proportions for $K = 3$, with individuals sorted by contact zone of origin and by major ancestry proportion.



Supplementary figure 2 Genomis PCA of the wild *P. axillaris* x *P. exserta* hybrids. A) First and second and B) first and third genomic PC of the analysed individuals. The colour represents the location where each individual comes from.



Supplementary figure 3 Pairwise phenotype correlation. Top right section displays Pearson's correlation value, bottom left displays Bonferroni-corrected *P* value.

Supplementary file 4: Excel table of phenotypic measurements, available on GitHub: htpps://github.com/Kuhlemeier-lab/genetics_petunia_hybrids.

References

Andrew, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Berardi, A.E., Esfeld, K., Jäggi, L., Mandel, T., Cannarozzi, G.M., and Kuhlemeier, C. (2021). Complex evolution of novel red floral color in *Petunia*. Plant Cell *33*, 2273–2295.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C.S., Bliek, M., Boersma, M.R., Borghi, L., Bruggmann, R., et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants *2*, 16074.

Bradshaw, H.D., and Schemske, D.W. (2003). Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature *426*, 176–178.

Bradshaw, H.D., Wilbert, S.M., Otto, K.G., and Schemske, D.W. (1995). Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). Nature *376*, 762–765.

Byers, K.J.R.P., Vela, J.P., Peng, F., Riffell, J.A., and Bradshaw, H.D. (2014). Floral volatile alleles can contribute to pollinator-mediated reproductive isolation in monkeyflowers (*Mimulus*). The Plant Journal *80*, 1031–1042.

Caballero-Villalobos, L., Silva-Arias, G.A., Turchetto, C., Giudicelli, G.C., Petzold, E., Bombarely, A., and Freitas, L.B. (2021). Neutral and adaptive genomic variation in hybrid zones of two ecologically diverged *Petunia* species (Solanaceae). Botanical Journal of the Linnean Society *196*, 100–122.

Carrió, E., and Güemes, J. (2014). The effectiveness of pre- and post-zygotic barriers in avoiding hybridization between two snapdragons (*Antirrhinum* L.: Plantaginaceae). Botanical Journal of the Linnean Society *176*, 159–172.

Chen, K.-Y., Cong, B., Wing, R., Vrebalov, J., and Tanksley, S.D. (2007). Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. Science *318*, 643–645.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*.

Dell'Olivo, A., and Kuhlemeier, C. (2013). Asymmetric effects of loss and gain of a floral trait on pollinator preference. Evolution *67*, 3023–3031.

Dell'Olivo, A., Hoballah, M.E., Gübitz, T., and Kuhlemeier, C. (2011). Isolation barriers between *Petunia axillaris* and *Petunia integrifolia* (Solanaceae). Evolution *65*, 1979–1991.

Edwards, M.B., Choi, G.P.T., Derieg, N.J., Min, Y., Diana, A.C., Hodges, S.A., Mahadevan, L., Kramer, E.M., and Ballerini, E.S. (2021). Genetic architecture of floral traits in bee- and hummingbird-pollinated sister species of *Aquilegia* (Columbine). Evolution *75*, 2197–2216.

Esfeld, K., Berardi, A.E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. (2018). Pseudogenization and resurrection of a speciation gene. Current Biology *28*, 3776-3786.e7.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology *14*, 2611–2620.

Fenster, C.B., Armbruster, W.S., Wilson, P., Dudash, M.R., and Thomson, J.D. (2004). Pollination syndromes and floral specialization. Annu. Rev. Ecol. Evol. Syst. *35*, 375–403.

Fishman, L., Kelly, A.J., and Willis, J.H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. Evolution *56*, 2138–2155.

Foll, M., and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics *180*, 977–993.

Fumagalli, M., Vieira, F.G., Korneliussen, T.S., Linderoth, T., Huerta-Sánchez, E., Albrechtsen, A., and Nielsen, R. (2013). Quantifying population genetic differentiation from next-generation sequencing data. Genetics *195*, 979–992.

Gao, L., and Rieseberg, L.H. (2020). While neither universally applicable nor practical operationally, the biological species concept continues to offer a compelling framework for studying species and speciation. National Science Review *7*, 1398–1400.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T., and Kuhlemeier, C. (2013). Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. Current Biology *23*, 873–877.

Hermann, K., Klahre, U., Venail, J., Brandenburg, A., and Kuhlemeier, C. (2015). The genetics of reproductive organ morphology in two *Petunia* species with contrasting pollination syndromes. Planta *241*, 1241–1254.

Hoballah, M.E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., and Kuhlemeier, C. (2007). Single gene–mediated shift in pollinator attraction in *Petunia*. The Plant Cell *19*, 779–790.

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinformatics *15*, 356.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Lindtke, D., and Buerkle, C.A. (2015). The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution *69*, 1987–2004.

Lorenz-Lemke, A.P., Mäder, G., Muschner, V.C., Stehmann, J.R., Bonatto, S.L., Salzano, F.M., and Freitas, L.B. (2006). Diversity and natural hybridization in a highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. Molecular Ecology *15*, 4487–4497.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Murray, M.G., and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research *8*, 4321–4326.

Nakazato, T., Rieseberg, L.H., and Wood, T.E. (2013). The genetic basis of speciation in the *Giliopsis* lineage of *Ipomopsis* (Polemoniaceae). Heredity *111*, 227–237.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. (2014). PopGenome: an efficient Swiss army knife for population genomic analyses in R. Mol Biol Evol *31*, 1929–1936.

Quattrocchio, F., Wing, J., Woude, K. van der, Souer, E., Vetten, N. de, Mol, J., and Koes, R. (1999). Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color. The Plant Cell *11*, 1433–1444.

R Core Team (2021). R: a language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Ramsey, J., Bradshaw, H.D., and Schemske, D.W. (2003). Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (phrymaceae). Evolution *57*, 1520–1534.

RStudio Team (2015). RStudio: integrated development for R. (RStudio, Inc., Boston, MA).

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature Methods *9*, 676–682.

Schnitzler, C.K., Turchetto, C., Teixeira, M.C., and Freitas, L.B. (2020). What could be the fate of secondary contact zones between closely related plant species? Genet. Mol. Biol. *43*, e20190271.

Segatto, A.L.A., Cazé, A.L.R., Turchetto, C., Klahre, U., Kuhlemeier, C., Bonatto, S.L., and Freitas, L.B. (2014). Nuclear and plastid markers reveal the persistence of genetic identity: A new perspective on the evolutionary history of *Petunia exserta*. Molecular Phylogenetics and Evolution *70*, 504–512.

Shang, Y., Venail, J., Mackay, S., Bailey, P.C., Schwinn, K.E., Jameson, P.E., Martin, C.R., and Davies, K.M. (2011). The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. New Phytologist *189*, 602–615.

Sheehan, H., Moser, M., Klahre, U., Esfeld, K., Dell'Olivo, A., Mandel, T., Metzger, S., Vandenbussche, M., Freitas, L., and Kuhlemeier, C. (2016). *MYB-FL* controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nature Genetics *48*, 159–166.

Skotte, L., Korneliussen, T.S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. Genetics *195*, 693–702.

Stehmann, J.R., Lorenz-Lemke, A.P., Freitas, L.B., and Semir, J. (2009). The genus *Petunia*. In Petunia, (New York, NY: Springer New York), pp. 1–28.

Teixeira, M.C., Turchetto, C., Hartke, S., Schnitzler, C.K., and Freitas, L.B. (2019). Morphological and genetic perspectives of hybridization in two contact zones of closely related species of *Petunia* (Solanaceae) in southern Brazil. Acta Botanica Brasilica *33*, 734–740.

Teixeira, M.C., Turchetto, C., Maestri, R., and Freitas, L.B. (2020). Morphological characterization of sympatric and allopatric populations of *Petunia axillaris* and *P. exserta* (Solanaceae). Botanical Journal of the Linnean Society *192*, 550–567.

Turchetto, C., Lima, J.S., Rodrigues, D.M., Bonatto, S.L., and Freitas, L.B. (2015a). Pollen dispersal and breeding structure in a hawkmoth-pollinated Pampa grasslands species *Petunia axillaris* (Solanaceae). Ann Bot *115*, 939–948.

Turchetto, C., Segatto, A.L.A., Beduschi, J., Bonatto, S.L., and Freitas, L.B. (2015b). Genetic differentiation and hybrid identification using microsatellite markers in closely related wild species. AoB PLANTS *7*, plv084.

Turchetto, C., Segatto, A.L.A., Silva-Arias, G.A., Beduschi, J., Kuhlemeier, C., Bonatto, S.L., and Freitas, L.B. (2019a). Contact zones and their consequences: hybridization between two ecologically isolated wild *Petunia* species. Bot J Linn Soc *190*, 421–435.

Turchetto, C., Schnitzler, C.K., and Freitas, L.B. (2019b). Species boundary and extensive hybridization and introgression in *Petunia*. Acta Botanica Brasilica *33*, 724–733.

Turchetto, C., Segatto, A.L.A., Lima, J.S., Hartke, S., Reck-Kortmann, M., Bonatto, S.L., and Freitas, L.B. (2021). So close, so far: spatial genetic structure and mating system in Petunia exserta, an endemic from a peculiar landscape in the Brazilian Pampa grasslands. Botanical Journal of the Linnean Society.

Watanabe, H., Ando, T., Tsukamoto, T., Hashimoto, G., and Marchesi, E. (2001). Cross-compatibility of *Petunia exserta* with other *Petunia* taxa. Journal of the Japanese Society for Horticultural Science *70*, 33–40.

White, R.H., Stevenson, R.D., Bennett, R.R., Cutler, D.E., and Haber, W.A. (1994). Wavelength discrimination and the role of ultraviolet vision in the feeding behavior of hawkmoths. Biotropica *26*, 427–435.

Wolf, P.G., Campbell, D.R., Waser, N.M., Sipes, S.D., Toler, T.R., and Archibald, J.K. (2001). Tests of pre- and postpollination barriers to hybridization between sympatric species of *Ipomopsis* (Polemoniaceae). American Journal of Botany *88*, 213–219.

Xu, S., Schlüter, P.M., Grossniklaus, U., and Schiestl, F.P. (2012). The genetic basis of pollinator adaptation in a sexually deceptive orchid. PLoS Genet *8*, e1002889.

Yarahmadov, T., Robinson, S., Hanemian, M., Pulver, V., and Kuhlemeier, C. (2020). Identification of transcription factors controlling floral morphology in wild *Petunia* species with contrasting pollination syndromes. The Plant Journal *104*, 289–301.

Yuan, Y.-W., Sagawa, J.M., Young, R.C., Christensen, B.J., and Bradshaw, H.D. (2013). Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics *194*, 255–263.

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nature Genetics *44*, 821–824.

Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet *9*, e1003264.

Chapter 3

Scent emission in the wild hawkmoth-pollinated *Petunia axillaris* is not mediated by *ABCG1*

Marta Binaghi, Mathilde Chopy, Cris Kuhlemeier

Institute of Plant Sciences, University of Bern, Bern 3013, Switzerland

Keywords: Scent transporter, Petunia, pollination syndromes, evolution.

Author contributions

C.K. conceived the study. M.B. designed the experiments, executed the bioinformatic analyses and produced the figures. M.C. conducted the experiments for figure 3.4. M.B. wrote the manuscript with the supervision of C.K.

Abstract

Most flowering plants rely on animals to transfer their pollen, and the evolution of sets of floral traits tailored to specific pollinators is thought to be a driving force in the rapid diversification of the angiosperms. The diversity in these sets of traits can be the cause of reproductive isolation between species. The evolutionary path that generated the differences in these traits is therefore of great interest for understanding evolution and speciation. In the genus *Petunia*, the shift from the bee-pollinated ancestor *P. inflata* to the hawkmoth-pollinated *P. axillaris* was accompanied by an increase in the number and quantity of volatiles produced. The molecular changes that determine these differences in scent production are known but the process of scent emission has only been described in *P. hybrida* cv Mitchell, where the transporter *ABCG1* ensures scent emission. This research cultivar derives from a series of crosses between *P. axillaris* and a bee-pollinated species and its scent profile, as well as its floral pigmentation and morphology are similar to *P. axillaris*. These similarities between *P. axillaris* and *P. hybrida* cv Mitchell suggest that scent emission has been conserved between them. In the wild *Petunia* species, the evolution of different scent profiles could have therefore been associated with changes in the transporter *ABCG1*. Here we compare the *ABCG1* transporters of four *Petunia* species and show that the scent transporter Ph*ABCG1* is most similar to the bee-pollinated ancestor of *P. hybrida* cv Mitchell. Expression data show that *P. axillaris* does not express *ABCG1*; but the characterisation of the *ABCG* subfamily in *P. axillaris* provides alternative candidates for this role. RNAseq data highlight one candidate gene with an expression pattern similar to *P. hybrida ABCG1*. Therefore, although they have similar scent profiles, the genes involved in scent emission are not the same in *P. axillaris* and *P. hybrida* cv Mitchell. *P. axillaris* probably uses a different transporter that is only distantly related to *ABCG1* to effect scent emission, suggesting that the scent pathway was not inherited conservatively from *P. axillaris* to the research cultivar *P. hybrida* cv Mitchell.

Introduction

Relevance of pollination syndromes for the study of plant speciation

The genetic basis of evolution has been described by numerous theoretical models that can be supported or contradicted by empirical results (Orr, 2005). The classical model of evolution presented by Fisher (Fisher, 1930) proposed that evolutionary change is due to the fixation of changes in many genes, each with a small effect. Consequently, speciation was thought to occur primarily in the absence of gene flow. More recently, empirical studies have demonstrated that mutations of major effect can be the drivers of adaptation and speciation, modifying the historical view of an evolutionary process based on infinite mutations of small effect (Chan et al., 2010; Manceau et al., 2010; Orr, 1992; Ueshima and Asami, 2003). New mathematical models have demonstrated that mutations of large effect can actually be beneficial in the initial stages of adaptation, and suggest that speciation is possible in the presence of gene flow (Orr, 1998; Orr and Coyne, 1992).

The majority of flowering plants recruit animals to transfer their pollen between individuals; floral traits evolve in response to pollinator availability, and these adaptive changes can result in reproductive isolation and speciation. Floral traits can be highly specific to one or few pollinators, and allow the classification of the flowers into distinct "pollination syndromes" (Fenster et al., 2004). Pollination syndromes are combinations of floral traits such as corolla pigmentation, scent and nectar production and morphology that attract and reward specific pollinators or pollinator groups. Shifts in pollination syndromes have been observed in many taxa and are thought to have driven the rapid diversification of flowering plants (Fenster et al., 2004; Grant, 1949; Mitchell et al., 2009). The specificity in the recruited pollinator restricts the mating possibilities of the plant and can therefore result in reproductive isolation even in sympatric conditions (Xu et al., 2011).

Petunia *evolved with its pollinators*

In the genus *Petunia*, shifts in pollination syndromes have happened several times. The genus is endemic to South America, and can be divided in two main clades based on the shape of the corolla. The ancestral short tube clade is more numerous and is characterised by a bee pollination syndrome presenting purple flowers and wide corolla tubes (Gübitz et al., 2009; Reck-Kortmann et al., 2014; Stehmann et al., 2009). *P. inflata* belongs to this clade and emits small amounts of volatiles during the daytime (Figure 3.1, Hoballah et al., 2005). The more recent long tube clade includes three different pollination syndromes (Figure 3.1 A): *P. axillaris* has white and UV absorbent flowers, emits large amounts of volatiles at dusk and is hawkmoth-pollinated (Ando et al., 2001; Hermann et al., 2013; Hoballah et al., 2005); *P. exserta* has bright red, UV-reflective flowers, and is hummingbird-pollinated (Griesbach et al., 1999; Hermann et al., 2013; Lorenz-Lemke et al., 2006; Sheehan et al., 2016); and *P. secreta* has purple and UV-reflective flowers and is bee-pollinated (Esfeld et al., 2018; Rodrigues et al., 2018; Stehmann and Semir, 2005).

The most recent phylogeny based on 300,000 SNPs supports an ancestral shift from the bee pollination of the short tube clade to the hawkmoth pollination presented by *P. axillaris* (Figure 3.1 A, Esfeld et al., 2018). This shift is characterised by the loss of purple colour and the gain of UV absorbance in the white *P. axillaris* flowers, two phenotypic changes caused by simple molecular modifications (Dell'Olivo and Kuhlemeier, 2013; Hoballah et al., 2007; Sheehan et al., 2016). The shift to the red flowers of *P. exserta* instead is due to changes in multiple regulatory and biosynthetic genes, contrasting to the subsequent simple re-acquisition of purple colour in *P. secreta* (Berardi et al., 2021; Esfeld et al., 2018). Different genetic architectures are therefore at the base of pigment changes in the long-tube clade of *Petunia*.

While geography is the major contributor to the separation of interfertile *Petunia* species, species of the long tube clade occasionally occur in sympatry and produce viable interspecific offspring (Lorenz-Lemke et al., 2006; Rodrigues et al.,

2018; Segatto et al., 2014; Turchetto et al., 2019). In this clade, pollinator preferences have been suggested as important reproductive isolation mechanisms, therefore making this clade a good system for the study of the genetic basis of reproductive isolation (Dell'Olivo et al., 2011).

In the long tube clade of Petunia*, scent production has changed several times; only a few mutations in key genes were necessary to cause these changes.*

Scent is an important signal between the plant and the environment. The volatile compounds emitted by the flowers can attract specific pollinators and promote floral visitation even when other attractive cues are present (e.g. floral colour, Klahre et al., 2011; Raguso and Willis, 2002, 2005). In many species, scent emission varies rhythmically during the day and correlates with the activity of the pollinators (Dötterl et al., 2005; Effmert et al., 2005; Raguso et al., 2003; Rodriguez-Saona et al., 2011). The presence, composition, and rhythm of emission of the floral bouquet determines which animals are attracted to the flowers and ultimately can result in reproductive isolation between plants that show differences in one or more of these scent characteristics (Byers et al., 2014; Dötterl et al., 2006; Klahre et al., 2011; Shuttleworth and Johnson, 2010; Vereecken et al., 2010; Waelti et al., 2008).

The scent pathway has been well documented in *Petunia hybrida* and most of the genes involved are known (Dudareva et al., 1996; Schlüter et al., 2011). In a few cases, the genetic modifications that caused phenotypic differences between wild species have been characterised. While the ancestral and bee-pollinated *P. inflata* produces only benzaldehyde, the moth-pollinated *P. axillaris* has gained in both the amount and in the diversity of volatiles produced; these changes are due to cis-acting polymorphisms that lead to the activation of two enzymes of the phenylpropanoid/benzenoid pathway (BPBT and BSMT), and to a general upregulation of the structural genes of the pathway (Figure 3.1 B, Amrad et al., 2016). The resultant change in the production of volatiles has been shown to influence attraction and visitation by hawkmoths (Klahre et al., 2011). In the hummingbird-

pollinated *P. exserta*, the loss of scent production is due to several null mutations in the coding sequence of the biosynthetic enzyme cinnamate-CoA ligase (CNL) and to the reduced expression of the MYB transcription factor ODORANT1 (ODO1, Amrad et al., 2016). Thus, the main changes in scent production that accompanied the evolution of pollination syndromes in the *Petunia* genus were achieved by means of few mutations in key genes involved in scent biosynthesis.

Scent emission

Scent production is only effective if the volatiles that are produced inside the cells are emitted into the environment. In the flowers, the main site of volatile production is the epidermis of the petal limb (Dudareva et al., 1996; Kolosova et al., 2001a), and the compounds have to cross at a minimum the plasma membrane, the cell wall, and the cuticle to be emitted. The volatiles can reach the environment by passive diffusion or active transport (Dudareva and Pichersky, 2006). Passive diffusion has been considered for a long time the main volatile emission route. However, it has been argued that to obtain the observed emission rates by passive diffusion, the concentration of volatiles in the cells would have to be high to the point of being cytotoxic (Widhalm et al., 2015). A system of active transport was therefore hypothesised.

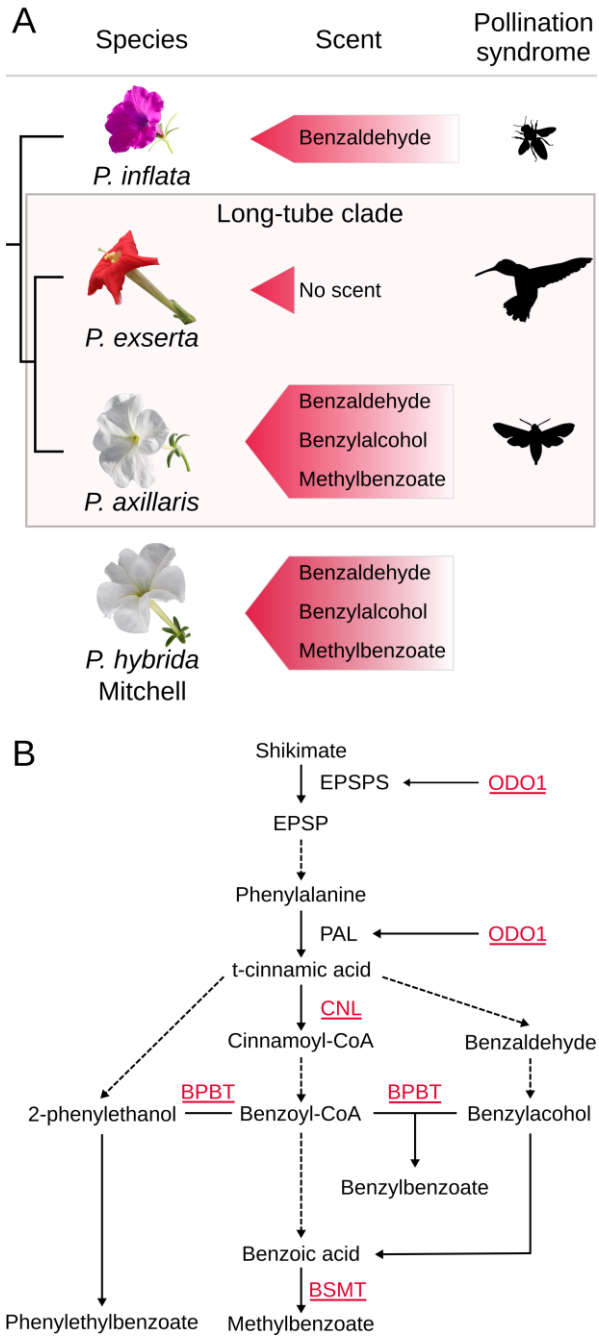


Figure 3.1 Scent evolution in *Petunia* species. A) Species of the genus *Petunia*, their pollinators, and the standard research cultivar *P. hybrida* cv Mitchell. The major scent compounds produced and emitted by each species are listed. The pollination syndromes are shown by animal silhouettes representing bee, hummingbird and hawkmoth. The phylogeny shown is according to Esfeld et al. 2018, and is based on 399,673 nucleotides. The long tube clade is boxed. B) Scent production-related proteins with roles in *Petunia* speciation. Red and underlined names are the enzymes and transcription factors that cause changes in the scent profile between the *Petunia* species. Dashed lines summarise multiple biosynthetic steps. BPBT Benzoyl-CoA:benzylalcohol/2-phenylethanol benzoyltransferase. BSMT Benzoic acid/salicylic acid carboxyl methyltransferase.

In P. hybrida*, scent emission is facilitated by an adenosine triphosphate-binding cassette (ABC) transporter.*

In plant cells, hydrophobic compounds such as waxes can be actively transported through the plasma membrane by ATP-binding cassette (ABC) transporters (Pighin et al., 2004). ABCG transporters constitute a numerous subfamily in plants (Mishra et al., 2019; Ofori et al., 2018); they are localised at the plasma membrane, and are involved in cuticle formation, defence mechanisms, hormone transport and seed germination (Bessire et al., 2011; Kang et al., 2010, 2015; Sasse et al., 2015, 2016). Their structure typically includes a nucleotide-binding domain (NBD) and a trans-membrane domain (TMD) organised as one or two pairs (Lefèvre and Boutry, 2018).

The flowers of the popular research cultivar *P. hybrida* cv Mitchell emit methylbenzoate and benzyl alcohol in a circadian rhythm which peaks at dusk (Verdonk et al., 2003). An analysis comparing transcript expression in flowers before and after opening at the peak of scent emission allowed Adebesin and colleagues (Adebesin et al., 2017) to identify an ABC transporter of the subfamily G, *ABCG1*, possibly involved in scent emission. The *P. hybrida* cv Mitchell ABCG1 is a half-size transporter, having only one trans-membrane and one nucleotide-binding domain. The protein is localized at the plasma membrane, and is hypothesised to work as a homodimer (Adebesin et al., 2017). The knockdown of the transporter by RNAi resulted in a reduced emission of volatiles and in their consequent accumulation inside the cells to the point that cytotoxic effects were visible (Adebesin et al., 2017). More recently, in a different *P. hybrida* line (P720), *ABCG1* transient overexpression was shown to cause an increase of the emission of volatiles from the abaxial side of the limb, further supporting the role of this transporter (Skaliter et al., 2021). Consistent with its role in scent emission, the gene is induced by the transcription factor ODO1 (Van Moerkercke et al., 2012), a major regulator of scent production in *Petunia* (Figure 3.1 B, Verdonk et al., 2005).

Is the role of ABCG1 in scent emission conserved in the evolutionary history of Petunia*?*

The identification of the scent transporter ABCG1 in *P. hybrida* cv Mitchell gives us the opportunity to investigate the role of this scent transporter in the wild *Petunia* species. The flowers of *P. hybrida* cv Mitchell are similar to those of *P. axillaris*: they have white petals and a narrow corolla tube and emit comparable bouquets of volatiles (Figure 3.1 A, Hoballah et al., 2005; Kolosova et al., 2001b; Negre et al., 2003; Verdonk et al., 2003). In *P. hybrida* cv Mitchell, the majority of the approximately 20,000 analysed genes can be assigned to *P. axillaris* (about 15,000), with only approximately 600 genes assigned to *P. inflata* (Bombarely et al., 2016). This indicates that the *P. inflata*-type parent makes only a minor contribution to the *P. hybrida* cv Mitchell gene space.

P. axillaris relies on scent emission to attract its nocturnal hawkmoth pollinator (Hoballah et al., 2005). We wanted to investigate whether, as part of its adaptation to the new pollinator, *P. axillaris* might not only have modified genes involved in scent production but also in scent transport. Based on the similarity of the scent profiles we hypothesised that *P. hybrida* cv Mitchell acquired its ABCG1 transporter from *P. axillaris* and that its function is conserved between these two species. Here we report that *P. hybrida* did not inherit its scent transporter from *P. axillaris* but from a short-tube bee-pollinated species. Moreover, we made the surprising discovery that wild *P. axillaris* does not express *ABCG1* and most likely employs a distantly related ABCG transporter.

Results

The protein sequence of ABCG1 *is generally conserved in wild* Petunia

The transporter ABCG1 is required for scent emission in *P. hybrida* cv Mitchell (Adebesin et al., 2017). The wild parental species of this cultivar are related to *P. axillaris* and *P. inflata*, two species that emit scent to attract their pollinators (Hoballah et al., 2005). We therefore hypothesised a role for ABCG1 as a scent transporter in the wild species. The closest homologues of Ph*ABCG1* were identified with a BLASTP search on the predicted protein sequences of the wild *Petunia* species, and validated with BLASTN on the genomes and transcriptomes (Supplementary file S1, Altschul et al., 1990). The protein sequences were then aligned with MUSCLE and a phylogenetic tree was constructed with PhyML (Supplementary file S2, Edgar, 2004; Guindon et al., 2010).

Among the three wild species, *P. hybrida* cv Mitchell has the most similar sequence to *P. inflata* (Figure 3.2). Compared to the *P. hybrida* cv Mitchell sequence, *P. inflata* has 12 amino acid changes, all outside of the annotated ABC transporter motifs. *P. axillaris* instead has 15 amino acid changes, plus one insertion and one deletion. The differences in protein sequence do not affect the canonical motifs of the nucleotide-binding domain in any of the species. In the C-term region, *P. axillaris* and *P. exserta* present one amino acid change in one transmembrane helix. These differences suggest that the proteins can be functional in all species. Note that evidence for the identification of the substrate specificity site are scarce, and that mutations in a specific site can affect transport of one substrate but not affect others (Lefèvre and Boutry, 2018). A phylogeny computed with PhyML suggests that the protein sequence of *P. hybrida* cv Mitchell is more closely related to *P. inflata* (branch support 0.70; Figure 3.2 B), than it is to the species of the long-tube clade. These results show that the protein sequence of *ABCG1* is generally well conserved in the wild species and that the *P. hybrida* cv Mitchell sequence is most similar to that of *P. inflata*. Thus, while *P. hybrida* cv

Mitchell is most similar to *P. axillaris* in many aspects, and particularly in the complexity and concentration of its scent bouquet, its scent transporter was probably inherited from a *P. inflata*-like parent.

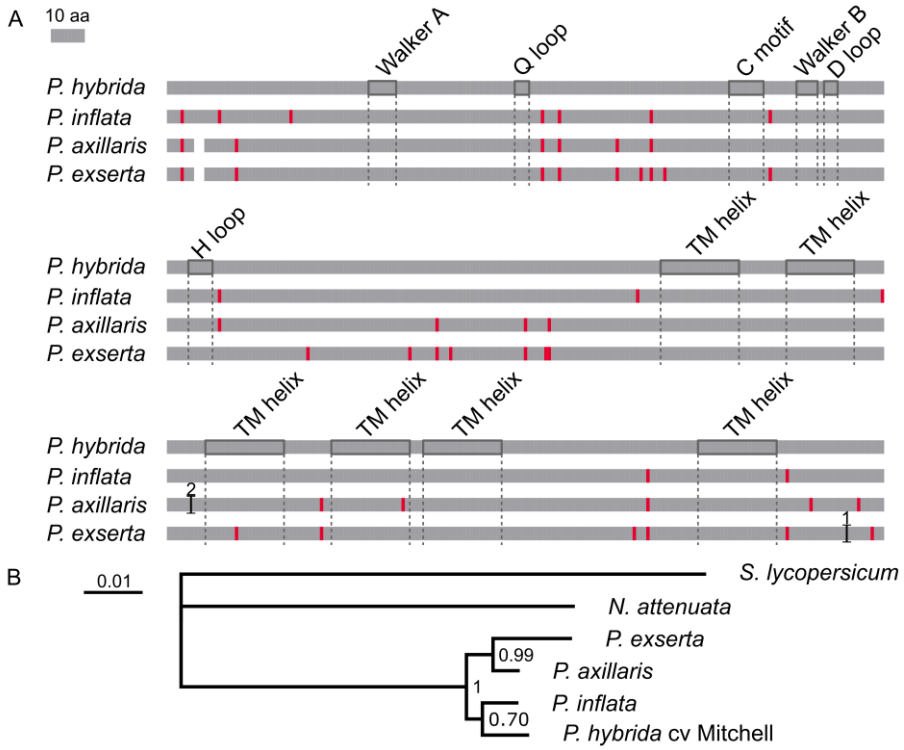


Figure 3.2 Comparison of the protein sequence of ABCG1 from the wild *Petunia* to the reference *P. hybrida* cv Mitchell. A) Graphic representation of the alignment. Grey colour indicates amino acid identity with the reference protein. Red indicates amino acid changes to the reference (the non-reference amino acid at any given position is always shared among non-reference sequences). White indicates deletions and the black "I" symbol shows insertions. The length of the insertion is indicated above the symbol. Dark grey frames indicate the position of the protein motifs. B) PhyML tree of protein sequences. 637 sites. Branch support is calculated using aLRT.

The scent transporter ABCG1 is not expressed in the scented P. axillaris

Starting at anthesis, *P. axillaris* flowers produce and emit scent in a circadian rhythm that peaks at dusk (Hoballah et al., 2005). This highly dynamic scent production is under the control of the transcription factor ODO1 (Verdonk et al., 2005). We therefore used RNAseq data to verify the presence of the *ABCG1* transcript in the petals of *P. axillaris* and to compare its level with that of *ODO1*. The

expression of *ODO1* is low at bud stage 4, but strongly increases in both tube and limb tissues of the open flowers, as expected (Figure 3.3). In contrast, *P. axillaris ABCG1* expression is negligible in the whole petal (tube and limb) at both bud and open flower stage.

ABCG1 expression at the peak of scent emission (ten minutes before dark) was measured with qRT-PCR to directly compare plants of *P. axillaris* with plants of *P. hybrida* cv Mitchell grown at the same conditions. The relative expression of *ABCG1* in *P. axillaris* is extremely low while it is high in *P. hybrida* cv Mitchell (Figure 3.4 C). We performed PCR and RT-PCR in two independent *P. axillaris* accessions (N and P) to confirm the qRT-PCR results (Figure 3.4 A and B). We conclude that the scent transporter *ABCG1* is not expressed in *P. axillaris*, and therefore a different transporter must be responsible for the emission of scent in this wild species.

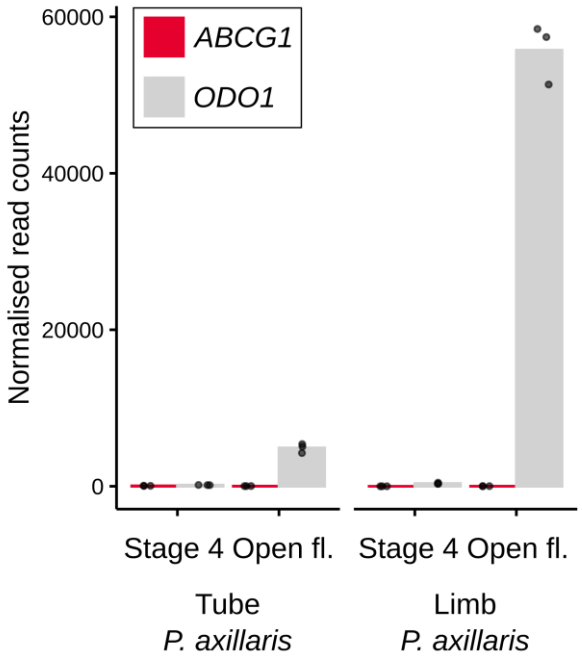


Figure 3.3 Expression of *ABCG1* and *ODO1* in *P. axillaris*. Normalised read counts of the transcripts of *ODO1* and *ABCG1* in *P. axillaris* limb and tube, at stage 4 and open flower. Bar height represents the mean and points represent single biological replicate values.

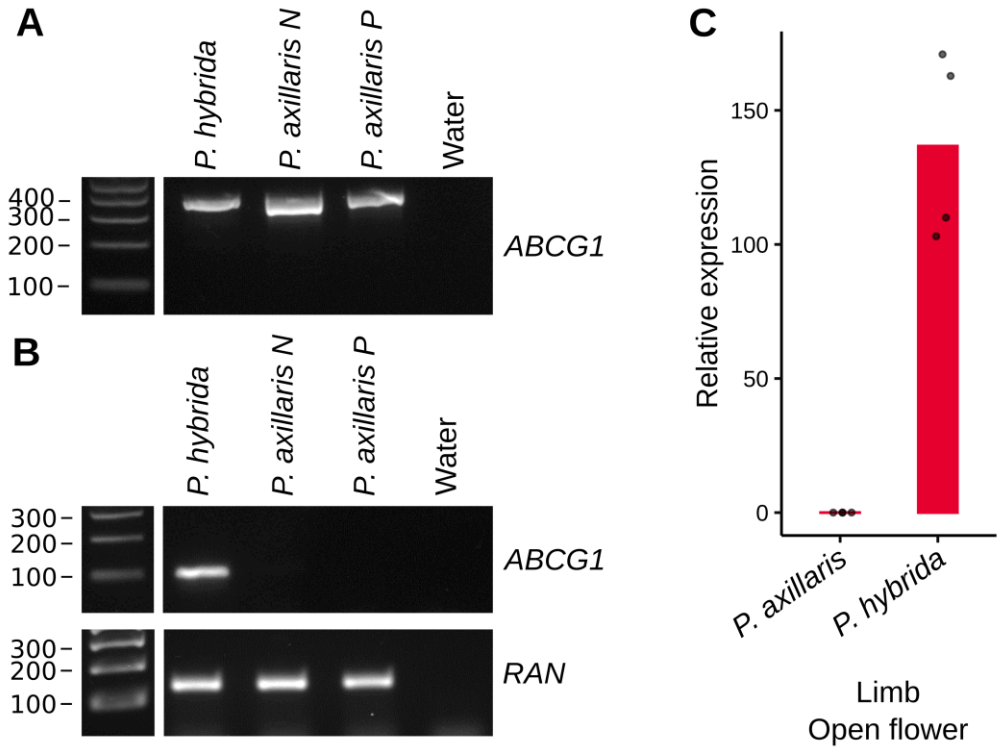


Figure 3.4 Expression of *ABCG1* in different *Petunia* species. A) PCR on genomic DNA from *P. hybrida* cv Mitchell, *P. axillaris* N and *P. axillaris* P, using *ABCG1* primers that amplify a 371 bp sequence. B) RT-PCR with gene-specific primers on RNA from open flower limbs of *P. hybrida* cv Mitchell, *P. axillaris* N and *P. axillaris* P. *ABCG1* primers amplify a 100 bp sequence. C) Relative expression of *ABCG1* in *P. axillaris* and *P. hybrida* cv Mitchell limb tissue at the peak of scent emission (ten minutes before dark), measured by qRT-PCR. Bar height represents the mean and points represent single biological replicate values.

The ABCG transporter Peaxi162Scf00060g00711 *is a candidate for scent emission in* P. axillaris

We hypothesise that another ABCG transporter may effect scent emission in *P. axillaris*. To find the scent transporter candidate in *P. axillaris* we identified all *ABCG*s of the species and analysed their expression patterns. We classified genes as *ABCG*s based on the following criteria: 1) similarity to one or more members of the *A. thaliana ABCG* subfamily, 2) presence of the ABC transporter signature motif (Pfam PF00005), and 3) presence of two or more transmembrane domains after the signature motif. Sixty-three genes were identified: 44 are half-size transporters (one nucleotide-binding and one trans-membrane domain), 16 are full-size

(two nucleotide-binding and two trans-membrane domains), and three have more than two pairs of domains (Figure 3.5).

We obtained the expression patterns of the 63 *ABCG*s in *P. axillaris* floral tissues using the RNAseq datasets already described (Supplementary file S4). Of the 63 *ABCG*s identified, 32 are expressed at the peak of scent emission (more than 10 normalised read counts, black font in Figure 3.5 A, listed in Table 3.1). To identify a potential scent transporter among the genes expressed we used two criteria: increased expression during the floral development between stage 4 (bud) and open flower limbs (log_2FC higher than one and adjusted *P* value lower than 0.05), and high expression at the peak of scent emission (more than 200 normalised read counts, Table 3.1). This selection identifies five candidate genes widely distributed in the cladogram (Figure 3.5 A). In the full-size clade, *Peaxi162Scf00423g00137* is extremely upregulated in the open flower limb (6.36 log_2FC) but its absolute expression only reaches 1013 normalised read counts at the peak of scent emission (Table 3.1). *Peaxi162Scf00938g00311* is not highly upregulated (1.22 log_2FC), but its expression at the peak of scent emission is the second highest in the dataset (3,719 normalised read counts). The third candidate in the full-size clade is modestly upregulated (2.42 log_2FC) and lowly expressed at the peak of scent emission (404 normalised read counts). In the half-size clade, *Peaxi162Scf00193g00724* shows a modest fold change (1.38 log_2FC) and low expression at the peak of scent emission (544 normalised read counts), and is highly expressed in the tube tissues, suggesting a possible specific role in the tube tissue (Figure 3.5). *Peaxi162Scf00060g00711* instead is 150 times more expressed in the open flower limb (7.25 log_2FC) and has the highest expression at the peak of scent emission. *Peaxi162Scf00060g00711* expression pattern is consistent with a role in scent emission and is similar to that of Ph*ABCG1* (Van Moerkercke et al., 2012). For these reasons we consider *Peaxi162Scf00060g00711* the best candidate for scent transport in *P. axillaris*.

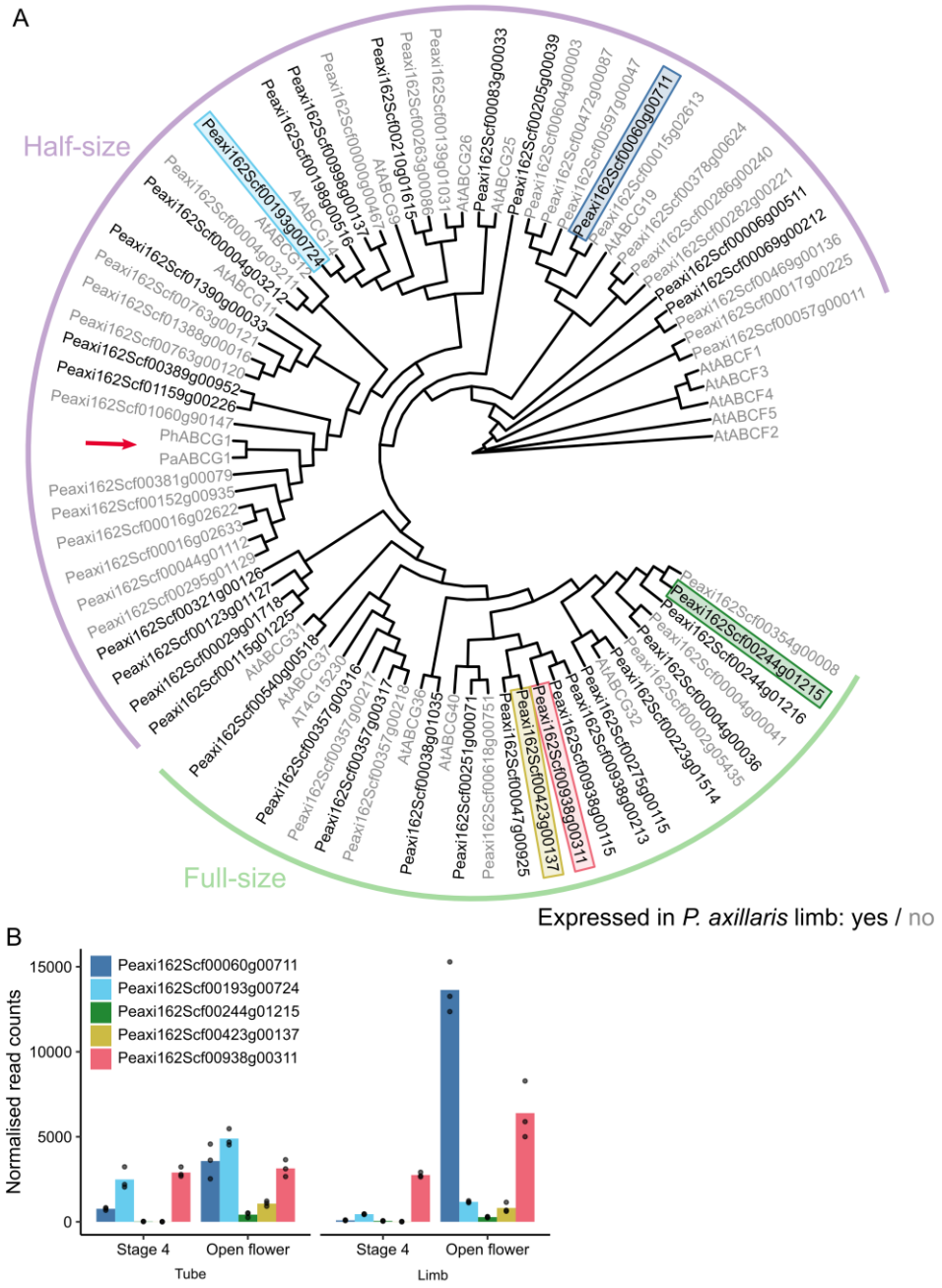


Figure 3.5 *ABCG* transcripts in *P. axillaris* limb and their expression. A) Cladogram of in silico translated transcript sequences of the *P. axillaris* ABCG subfamily. The tree includes *P. hybrida* cv Mitchell *ABCG1*, several *A. thaliana ABCG*s, and the *A. thaliana ABCF* subfamily as outgroup. Transcripts that are expressed in *P. axillaris* limb at the peak of scent emission (more than 10 normalised read counts) are written in black, transcripts not expressed or from a different species are written in grey. Green and pink curves indicate half- and full-size transporters. Coloured backgrounds indicate correspondence to transcripts in panel B. B) Expression in *P. axillaris* floral tissues of the five most highly expressed *ABCG*s in *P. axillaris* limb and tube at stage 4 and open flower. Bar height represents the mean and points represent single biological replicate values.

Table 3.1 Expression of *ABCG* transcripts in *P. axillaris* limb tissue[a]

	Normalised read count, mean (SD)			DE stage 4 VS open fl.[b]	
Gene ID	Emission peak	Stage 4	Open flower	log2(FC)	*P* adj
Peaxi162Scf00060g00711	8032 (1804)	88 (22)	13636 (1499)	7.25	0
Peaxi162Scf00938g00311	3719 (713)	2753 (145)	6392 (1696)	1.22	0
Peaxi162Scf00004g03212	1177 (735)	3647 (365)	2864 (264)	-0.35	0.2
Peaxi162Scf00004g00036	1068 (306)	1702 (288)	2165 (137)	0.35	0.07
Peaxi162Scf00423g00137	1013 (506)	11 (9)	815 (290)	6.36	0
Peaxi162Scf01390g00033	728 (111)	10242 (572)	1090 (90)	-3.23	0
Peaxi162Scf00205g00039	559 (43)	1253 (45)	923 (81)	-0.44	0
Peaxi162Scf00193g00724	544 (290)	449 (21)	1172 (59)	1.38	0
Peaxi162Scf00244g01215	404 (17)	51 (15)	276 (37)	2.42	0
Peaxi162Scf00321g00126	372 (68)	728 (194)	750 (20)	0.03	0.89
Peaxi162Scf00210g01615	302 (78)	4483 (707)	540 (208)	-3.05	0
Peaxi162Scf00223g01514	254 (41)	8910 (1083)	605 (186)	-3.88	0
Peaxi162Scf00198g00516	173 (27)	2 (2)	134 (18)	5.56	0
Peaxi162Scf00389g00952	133 (22)	577 (17)	275 (3)	-1.06	0
Peaxi162Scf00083g00033	131 (14)	2503 (321)	257 (19)	-3.29	0
Peaxi162Scf00038g01035	106 (114)	789 (348)	77 (43)	-3.36	0
Peaxi162Scf00998g00137	99 (108)	69 (94)	48 (21)	-0.53	0.65
Peaxi162Scf00251g00071	88 (19)	56 (17)	141 (39)	1.34	0.02
Peaxi162Scf01159g00226	87 (13)	573 (50)	129 (15)	-2.15	0
Peaxi162Scf00123g01127	79 (88)	54 (72)	17 (19)	-1.72	0.16
Peaxi162Scf00938g00213	71 (36)	88 (16)	206 (53)	1.18	0
Peaxi162Scf00357g00317	48 (21)	1 (1)	56 (21)	6.05	0
Peaxi162Scf00244g01216	48 (20)	4166 (481)	104 (28)	-5.32	0
Peaxi162Scf00006g00511	31 (6)	184 (27)	91 (9)	-1	0
Peaxi162Scf00938g00115	30 (4)	36 (10)	73 (17)	1.11	0
Peaxi162Scf00115g01225	28 (40)	27 (29)	10 (9)	-1.44	0.1
Peaxi162Scf00029g01718	21 (24)	19 (32)	4 (4)	-2.47	0.19
Peaxi162Scf00540g00518	20 (6)	81 (55)	87 (25)	0.14	0.85
Peaxi162Scf00047g00925	16 (5)	98 (30)	19 (6)	-2.35	0
Peaxi162Scf00357g00316	15 (11)	1 (2)	7 (2)	3.05	0.02
Peaxi162Scf00275g00115	15 (5)	1020 (105)	64 (25)	-3.99	0
Peaxi162Scf00069g00212	11 (4)	0 (0)	26 (5)	6.62	0

[a] Only *ABCG* transcripts with expression higher than 10 normalised read counts at the moment of scent emission peak (ten minutes before dark) are shown. The read counts are the result of two different experiments: one at the peak of scent emission, one at stage 4 and open flower. Full table in supplementary file S4.

[b] Differential expression (DE) analysis between stage 4 and open flower limbs. Log_2(fold change) and adjusted *P* value.

Discussion

P. hybrida cv Mitchell is a research cultivar whose origin dates back to the 1980s, when Adele Z. Mitchell produced a double haploid line derived from a cross between *P. axillaris* × *P. hybrida* cv Rose du Ciel followed by one backcross with *P. axillaris* (Mitchell et al., 1980). Since then, the cultivar has been maintained by selfing and has been used to investigate different research topics, including scent biosynthesis and emission. Its scent scent profile is similar to that of *P. axillaris* and is characterised by high emission of methylbenzoate and lower emission of benzaldehyde (Boatright et al., 2004; Dell'Olivo and Kuhlemeier, 2013; Verdonk et al., 2003). The recent discovery of a scent transporter in this cultivar, ABCG1, has motivated us to investigate scent transport in the scented wild species of *Petunia* (Adebesin et al., 2017).

The protein sequence of the *P. hybrida* cv Mitchell transporter ABCG1 is closely related to those of the wild species, and shows the highest similarity to *P. inflata*, a lowly scented species of the short-tube clade (Figure 3.2). Thus we assume that *P. hybrida* cv Mitchell ABCG1 was contributed by the *P. inflata*-type parent. The transcripts of *P. inflata*-origin represent about 3% of the *P. hybrida* cv Mitchell transcriptome, with the remaining transcripts originating from its other ancestor, *P. axillaris* (Bombarely et al., 2016). The presence of a *P. inflata*-like transporter in this research cultivar is therefore remarkable, and could have been the result of linked selection during the development of the market cultivar Rose du Ciel. In commercial crosses, *P. axillaris* and related white cultivars are often used to introgress colour diversity from other backgrounds. The scent transporter gene could have been co-selected if it were genetically linked to a colour gene.

In the wild species *P. axillaris* our results clearly show that the homolog of *P. hybrida ABCG1* is not expressed (Figure 3.3 and 3.4). The wild *P. axillaris* produces an important amount of volatiles to attract its pollinator (Klahre et al., 2011). Therefore we expected to see the scent transporter *ABCG1* expressed in the floral tissue at the moment of high scent emission. Contrary to our expectations,

we could not find the transcript of this gene in such tissues. Given the importance of scent for this wild species, we hypothesised that another gene is responsible for the emission of scent. Our search for *ABCG* candidates for scent transport highlighted five genes with a compatible pattern, one of which is the most highly expressed *ABCG* and is extremely upregulated during floral development (Table 3.1). Even though none of the five genes is closely related to *ABCG1* (Figure 3.5), we couldn't find a closer candidate with a compatible expression pattern. These results suggest that scent emission in the wild *P. axillaris* is controlled by a different transporter or that scent emission in this species takes a different route. The candidate genes can be further investigated to identify the elusive scent transporter of *P. axillaris*. The results of future experiments could inform us on the evolutionary pattern of the scent transporter gene in relation to scent production evolution, indicating if the changes of scent production were closely associated to scent emission changes, and how these two sections of the scent pathway depend on each other.

Materials and methods

Plant material and growth conditions

Petunia axillaris N comes from the Rostock Botanical Garden (Germany) and is registered in the Amsterdam collection under the designation of *P. axillaris* S26. *Petunia hybrida* cv Mitchell (also known as *P. hybrida* W115) was provided by D. Reinhardt (University of Fribourg, Switzerland). *P. hybrida* cv Mitchell is a double haploid line derived from a cross between *P. axillaris* × *P. hybrida* cv Rose du Ciel followed by one backcross with *P. axillaris* (Mitchell et al., 1980) and has been propagated by selfing since the 1980s. Plants for PCR, RT-PCR and qRT-PCR experiments were grown in a growth chamber under a light:dark regime of 15:9, at 22:17°C at 60-80% relative humidity, in commercial soil (70% Klasman substrate, 15% Seramis clay granules, 15% quartz sand), and fertilized once a week.

Retrieval of ABCG1 *sequences in wild* Petunia *and sequence comparison*

The sequences of Ph*ABCG1* were obtained from GenBank accession JQ088099.1. The genomic sequence was used as a query in a BLASTN search (default parameters, version 2.9.0, Altschul et al., 1990) on a local BLAST server (sequenceserver version 1.0.8, Priyam et al., 2019) against the genomes of *P. axillaris* (version 1.6.2 available on Sol Genomics Network, https://solgenomics.net/organism/Petunia_axillaris/genome), *P. inflata* (version 1.0.1 available on Sol Genomics Network https://solgenomics.net/organism/Petunia_inflata/genome), *P. exserta* (in preparation). The annotations were confirmed and corrected when needed by RNAseq reads alignment. Transcripts and their in silico translated sequences are available in Supplementary file S1. Sequence alignments and tree building were performed in the interface provided by Unipro UGENE (version 37.0, Okonechnikov et al., 2012). The annotated coding sequences were translated in silico and the resulting protein sequences were aligned with MUSCLE v3.8.31 (default parameters, Ed-

gar, 2004). The *Petunia ABCG1* tree was built with PhyML version 20120412 with fast likelihood-based method for branch support (Guindon et al., 2010). Raw alignment of *Petunia ABCG1* sequences is available in Supplementary file S2.

RNAseq expression of ABCG *genes*

RNAseq expression data were obtained from previous experiments performed in the group. Samples included *P. axillaris* petals (tube and limb separately) at bud stage 4 (22-30 mm in length) and open flower (one day post-anthesis) collected during daytime. One additional experiment included *P. axillaris* limbs collect at the peak of scent emission (10 minutes before dark, on day 1 post-anthesis). RNA extraction, library preparation and sequencing settings as described in Berardi et al., 2021, limb data; Yarahmadov et al., 2020, tube data; and Amrad et al., 2016, limb data collected 10 minutes before dark. Raw reads are available under NCBI SRA accessions PRJNA720957, PRJNA674380, PRJNA724705 and PRJNA344710. Raw reads data were quality controlled with FastQC (Andrew, 2010). Trimming was performed with Trimmomatic (Bolger et al., 2014). Alignment performed with STAR (Dobin et al., 2013) on the reference genome of *P. axillaris* version 1.6.2 available on Sol Genomics Network (https://solgenomics.net/organism/Petunia_axillaris/genome) and read counts obtained with featureCounts (Liao et al., 2014), using the gene annotations associated to the reference genome. Differential expression analysis was performed with DeSeq2 (Love et al., 2014) in R (R Core Team, 2021), and plots were built with the R library ggplot2 (Wickham, 2016). Details on software versions and full scripts of this analysis are available at https://github.com/Kuhlemeier-lab/scent_transporter. Resource-demanding computations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

qRT-PCR, PCR and RT-PCR of ABCG1

qRT-PCRs were performed on flower limbs at one day post-anthesis for the species *P. axillaris* and *P. hybrida* cv Mitchell. One limb was collected for each of the four biological replicates and immediately frozen in liquid nitrogen. RNA extrac-

tion was performed with innuPREP RNA Mini Kit (Analytik Jena). DNase treatment was performed with Amplification Grade DNase I (Sigma-Aldrich). RNA was quantified using NanoDrop and 1µg was used for subsequent reverse transcription with qScriber cDNA Synthesis Kit (HighQu). PCR reactions were performed in an optical 384-well plate in the QuantStudioTM 5 Flex Real-Time PCR System (Applied Biosystems), using ORA™ SEE qPCR Green ROX L Mix, 2X, in a final volume of 10 µl, according to the manufacturer's instructions. The following standard thermal profile was used for all PCR reactions: 95 °C for 2 min, 40 cycles of 95 °C for 5 s, and 60 °C for 20 s. Data were analysed using the QuantStudio™ Design & Analysis Software (Applied Biosystems). Primer pairs were tested for efficiency calculation and specificity. Primer sequences are available in Supplementary file S3. The relative expression of *ABCG1* was calculated as the mean of the expression relative to the two housekeeping genes *RAN1* and *ACTIN11* calculated with the ΔΔCt method. Specific primer pairs were designed to amplify the products (Supplementary file S3).

For RT-PCR, the steps from RNA extraction to cDNA synthesis were performed as described before for qRT-PCR. A PCR was performed on the cDNA using GoTaq polymerase in a final volume of 20 µl. The following thermocycling conditions were used for RT-PCR reactions: 94 °C for 5 min, 30 cycles of 94 °C for 15 s, 55°C for 15 s and 72 °C 1min per kb. PCR products were separated by electrophoresis on a 1.5% agarose gel. Primers are listed in Supplementary file S3.

For the PCR on genomic DNA, extraction of DNA from young tissue was performed according to Edwards et al., 1991, and a PCR was performed to verify the efficiency of the primer pairs used for the RT-PCR using GoTaq polymerase in a final volume of 20 µl. PCR products were separated by electrophoresis on a 1.5% agarose gel. Primers are listed in Supplementary file S3.

ABCG *subfamily characterisation in* P. axillaris

To identify *ABCG* genes in the *P. axillaris* genome, the complete list of *A. thaliana ABCG* genes was downloaded from TAIR, download section, file

gene_families_sep_29_09_update.txt (The Arabidopsis Information Resource, on www.arabidopsis.org, accessed on 28/05/2020) and searched for loci annotated as *WBC* or *PDR*. The corresponding transcript sequences were obtained from TAIR10 genome release with Samtools (Danecek et al., 2021). A local BLAST database was compiled for the *P. axillaris* annotated transcripts and genome and used as search spaces for two BLAST searches (see GitHub repository for complete commands). Regions identified by the genome search but absent from the transcript search were extracted and used for ab initio gene prediction with AUGUSTUS (http://bioinf.uni-greifswald.de/augustus/, Keller et al., 2011); putative genes with transcripts longer than 1200 bases were kept. The four putative genes so discovered were added to the *P. axillaris* annotation file. InterProScan (Jones et al., 2014) was used to predict protein motifs and TMHMM to predict trans-membrane helices (http://www.cbs.dtu.dk/services/TMHMM/, Krogh et al., 2001; Sonnhammer et al., 1998). The transcripts whose predicted motifs included one or more modules composed of one ABC signature motif (Pfam PF00005) followed by two or more trans-membrane helices were considered as members of the *P. axillaris ABCG* subfamily and analysed further. The RNAseq datasets described above were queried for the *ABCG* transcript expression in tube and limb at stage 4 bud and open flower, and fold change in limb tissue between bud stage 4 and open flower was computed and reported with corresponding Bonferroni-corrected *P* values. To be sure of the expression of the genes at peak of scent emission, when Ph*ABCG1* is most expressed in *P. hybrida* cv Mitchell, the RNAseq dataset of *P. axillaris* limb tissues collected 10 minutes before dark was queried. The *ABCG* phylogenetic tree was built by aligning the ABC signature motif PF00005 of each in silico translated transcript. If more than one PF00005 motif was present, their sequences were concatenated. Five protein sequences (only the PF00005 motif) of the *A. thaliana* ABCF subfamily were used as an outgroup, according to the ABC subfamily phylogeny proposed by (Xiong et al., 2015). Alignment was performed in UGENE, using MUSCLE default parameters. Tree was computed with PhyML with fast likelihood-based method for branch support; visualisation and annotation were performed with the ggtree library in R (Yu, 2020; Yu et al., 2017). Steps for

sequence manipulation, gene set overlap, plot production, tree annotation and similar were performed in bash and R. Detailed scripts available on GitHub.

Data and code availability

Raw reads are available on NCBI SRA, under accessions PRJNA720957, PRJNA674380, PRJNA724705 and PRJNA344710. The code used for the analyses and the specific versions of the software used are available on GitHub at https://github.com/Kuhlemeier-lab/scent_transporter.

Supplemental data

Supplementary files are available on GitHub, at https://github.com/Kuhlemeier-lab/scent_transporter.

Supplementary file S1. Sequences of wild Petunia ABCG1s (S1_ABCG1_sequences_petunia.zip)

Supplementary file S2. Alignment of the Petunia ABCG1 (S2_ABCG1_petunia_prot.phy)

Supplementary file S3. Primer sequences for the PCR, RT-PCR and qRT-PCR (S3_primers_ABCG1.txt)

Supplementary file S4. RNAseq normalised read counts of the ABCG of P. axillaris (S4_abcg_paxi_rnaseq.xlsx)

Acknowledgements

We would like to thank Avichai Amrad for providing the initial spark of this project. We thank Gina Cannarozzi for her expert supervision on phylogenetic analyses, and Andrea Berardi for advice and help with scent measurement. We thank Jasmin Sekulowski and Christopher Ball for expert care of our plants, and Didier Reinhardt for providing *P. hybrida* cv Mitchell seeds.

References

Adebesin, F., Widhalm, J.R., Boachon, B., Lefèvre, F., Pierman, B., Lynch, J.H., Alam, I., Junqueira, B., Benke, R., Ray, S., et al. (2017). Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. Science *356*, 1386–1388.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403–410.

Amrad, A., Moser, M., Mandel, T., de Vries, M., Schuurink, R.C., Freitas, L., and Kuhlemeier, C. (2016). Gain and loss of floral scent production through changes in structural genes during pollinator-mediated speciation. Current Biology *26*, 3303–3312.

Ando, T., Nomura, M., Tsukahara, J., Watanabe, H., Kokubun, H., Tsukamoto, T., Hashimoto, G., Marchesi, E., and Kitching, I.J. (2001). Reproductive isolation in a native population of *Petunia sensu* Jussieu (Solanaceae). Annals of Botany *88*, 403–413.

Andrew, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Berardi, A.E., Esfeld, K., Jäggi, L., Mandel, T., Cannarozzi, G.M., and Kuhlemeier, C. (2021). Complex evolution of novel red floral color in *Petunia*. Plant Cell *33*, 2273–2295.

Bessire, M., Borel, S., Fabre, G., Carraça, L., Efremova, N., Yephremov, A., Cao, Y., Jetter, R., Jacquat, A.-C., Métraux, J.-P., et al. (2011). A member of the PLEIOTROPIC DRUG RESISTANCE family of ATP binding cassette transporters is required for the formation of a functional cuticle in *Arabidopsis*. The Plant Cell *23*, 1958–1970.

Boatright, J., Negre, F., Chen, X., Kish, C.M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D., and Dudareva, N. (2004). Understanding in vivo benzenoid metabolism in *Petunia* petal tissue. Plant Physiology *135*, 1993–2011.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C.S., Bliek, M., Boersma, M.R., Borghi, L., Bruggmann, R., et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants *2*, 16074.

Byers, K.J.R.P., Bradshaw, H.D., and Riffell, J.A. (2014). Three floral volatiles contribute to differential pollinator attraction in monkeyflowers (*Mimulus*). Journal of Experimental Biology *217*, 614–623.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. Science *327*, 302–305.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*.

Dell'Olivo, A., and Kuhlemeier, C. (2013). Asymmetric effects of loss and gain of a floral trait on pollinator preference. Evolution *67*, 3023–3031.

Dell'Olivo, A., Hoballah, M.E., Gübitz, T., and Kuhlemeier, C. (2011). Isolation barriers between *Petunia axillaris* and *Petunia integrifolia* (Solanaceae). Evolution *65*, 1979–1991.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dötterl, S., Wolfe, L.M., and Jürgens, A. (2005). Qualitative and quantitative analyses of flower scent in *Silene latifolia*. Phytochemistry *66*, 203–213.

Dötterl, S., Jürgens, A., Seifert, K., Laube, T., Weissbecker, B., and Schütz, S. (2006). Nursery pollination by a moth in *Silene latifolia*: the role of odours in eliciting antennal and behavioural responses. New Phytol *169*, 707–718.

Dudareva, N., and Pichersky, E. (2006). Biology of floral scent (CRC Press).

Dudareva, N., Cseke, L., Blanc, V.M., and Pichersky, E. (1996). Evolution of floral scent in *Clarkia*: novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. Plant Cell *8*, 1137–1148.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792–1797.

Effmert, U., Große, J., Röse, U.S.R., Ehrig, F., Kägi, R., and Piechulla, B. (2005). Volatile composition, emission pattern, and localization of floral scent emission in *Mirabilis jalapa* (Nyctaginaceae). Am J Bot *92*, 2–12.

Esfeld, K., Berardi, A.E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. (2018). Pseudogenization and resurrection of a speciation gene. Current Biology *28*, 3776-3786.e7.

Fenster, C.B., Armbruster, W.S., Wilson, P., Dudash, M.R., and Thomson, J.D. (2004). Pollination syndromes and floral specialization. Annu. Rev. Ecol. Evol. Syst. *35*, 375–403.

Fisher, R.A. (1930). The genetical theory of natural selection (Oxford, England: Clarendon Press).

Grant, V. (1949). Pollination systems as isolating mechanisms in angiosperms. Evolution *3*, 82–97.

Griesbach, R.J., Stehmann, J.R., and Meyer, F. (1999). Anthocyanins in the "red" flowers of *Petunia exserta*. Phytochemistry *51*, 525–528.

Gübitz, T., Hoballah, M.E., Dell'Olivo, A., and Kuhlemeier, C. (2009). Petunia as a model system for the genetics and evolution of pollination syndromes. In Petunia, T. Gerats, and J. Strommer, eds. (New York, NY: Springer New York), pp. 29–50.

Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol *59*, 307–321.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T., and Kuhlemeier, C. (2013). Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. Current Biology *23*, 873–877.

Hoballah, M.E., Stuurman, J., Turlings, T.C.J., Guerin, P.M., Connétable, S., and Kuhlemeier, C. (2005). The composition and timing of flower odour emission by wild *Petunia axillaris* coincide with the antennal perception and nocturnal activity of the pollinator *Manduca sexta*. Planta *222*, 141–150.

Hoballah, M.E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., and Kuhlemeier, C. (2007). Single gene–mediated shift in pollinator attraction in *Petunia*. The Plant Cell *19*, 779–790.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236–1240.

Kang, J., Hwang, J.U., Lee, M., Kim, Y.Y., Assmann, S.M., Martinoia, E., and Lee, Y. (2010). PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. Proceedings of the National Academy of Sciences of the United States of America (PNAS) *107*, 2355–2360.

Kang, J., Yim, S., Choi, H., Kim, A., Lee, K.P., Lopez-Molina, L., Martinoia, E., and Lee, Y. (2015). Abscisic acid transporters cooperate to control seed germination. Nature Communications *6*, 1–10.

Keller, O., Kollmar, M., Stanke, M., and Waack, S. (2011). A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics *27*, 757–763.

Klahre, U., Gurba, A., Hermann, K., Saxenhofer, M., Bossolini, E., Guerin, P.M., and Kuhlemeier, C. (2011). Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. Current Biology *21*, 730–739.

Kolosova, N., Sherman, D., Karlson, D., and Dudareva, N. (2001a). Cellular and subcellular localization of S-adenosyl-L-methionine:benzoic acid carboxyl methyltransferase, the enzyme responsible for biosynthesis of the volatile ester methylbenzoate in snapdragon flowers. Plant Physiol *126*, 956–964.

Kolosova, N., Gorenstein, N., Kish, C.M., and Dudareva, N. (2001b). Regulation of circadian methyl benzoate emission in diurnally and nocturnally emitting plants. The Plant Cell *13*, 2333–2347.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. Journal of Molecular Biology *305*, 567–580.

Lefèvre, F., and Boutry, M. (2018). Towards identification of the substrates of ATP-binding cassette transporters. Plant Physiol *178*, 18–39.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Lorenz-Lemke, A.P., Mäder, G., Muschner, V.C., Stehmann, J.R., Bonatto, S.L., Salzano, F.M., and Freitas, L.B. (2006). Diversity and natural hybridization in a highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. Molecular Ecology *15*, 4487–4497.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550.

Manceau, M., Domingues, V.S., Linnen, C.R., Rosenblum, E.B., and Hoekstra, H.E. (2010). Convergence in pigmentation at multiple levels: mutations, genes and function. Philosophical Transactions of the Royal Society B: Biological Sciences *365*, 2439–2450.

Mishra, A.K., Choi, J., Rabbee, M.F., and Baek, K.-H. (2019). In silico genome-wide analysis of the ATP-binding cassette transporter gene family in soybean (*Glycine max* L.) and their expression profiling. Biomed Res Int *2019*.

Mitchell, A.Z., Hanson, M.R., Skvirsky, R.C., and Ausubel, F.M. (1980). Anther culture of *Petunia*: genotypes with high frequency of callus, root, or plantlet formation. Zeitschrift Für Pflanzenphysiologie *100*, 131–145.

Mitchell, R.J., Irwin, R.E., Flanagan, R.J., and Karron, J.D. (2009). Ecology and evolution of plant–pollinator interactions. Ann Bot *103*, 1355–1363.

Negre, F., Kish, C.M., Boatright, J., Underwood, B., Shibuya, K., Wagner, C., Clark, D.G., and Dudareva, N. (2003). Regulation of methylbenzoate emission after pollination in snapdragon and petunia flowers. Plant Cell *15*, 2992–3006.

Ofori, P.A., Mizuno, A., Suzuki, M., Martinoia, E., Reuscher, S., Aoki, K., Shibata, D., Otagaki, S., Matsumoto, S., and Shiratake, K. (2018). Genome-wide analysis of ATP binding cassette (ABC) transporters in tomato. PLoS One *13*.

Okonechnikov, K., Golosova, O., Fursov, M., and the UGENE team (2012). Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics *28*, 1166–1167.

Orr, H.A. (1992). Mapping and characterization of a "speciation gene" in *Drosophila*. Genet Res *59*, 73–80.

Orr, H.A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution *52*, 935–949.

Orr, H.A. (2005). The genetic theory of adaptation: a brief history. Nature Reviews Genetics *6*, 119–127.

Orr, H.A., and Coyne, J.A. (1992). The genetics of adaptation: a reassessment. Am Nat *140*, 725–742.

Pighin, J.A., Zheng, H., Balakshin, L.J., Goodman, I.P., Western, T.L., Jetter, R., Kunst, L., and Samuels, A.L. (2004). Plant cuticular lipid export requires an ABC transporter. Science *306*, 702–704.

Priyam, A., Woodcroft, B.J., Rai, V., Moghul, I., Munagala, A., Ter, F., Chowdhary, H., Pieniak, I., Maynard, L.J., Gibbins, M.A., et al. (2019). Sequenceserver: a modern graphical user interface for custom BLAST databases. Molecular Biology and Evolution *36*, 2922–2924.

R Core Team (2021). R: a language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Raguso, R.A., and Willis, M.A. (2002). Synergy between visual and olfactory cues in nectar feeding by naïve hawkmoths, *Manduca sexta*. Animal Behaviour *64*, 685–695.

Raguso, R.A., and Willis, M.A. (2005). Synergy between visual and olfactory cues in nectar feeding by wild hawkmoths, *Manduca sexta*. Animal Behaviour *69*, 407–418.

Raguso, R.A., Levin, R.A., Foose, S.E., Holmberg, M.W., and McDade, L.A. (2003). Fragrance chemistry, nocturnal rhythms and pollination "syndromes" in Nicotiana. Phytochemistry *63*, 265–284.

Reck-Kortmann, M., Silva-Arias, G.A., Segatto, A.L.A., Mäder, G., Bonatto, S.L., and de Freitas, L.B. (2014). Multilocus phylogeny reconstruction: new insights into the evolutionary history of the genus *Petunia*. Molecular Phylogenetics and Evolution *81*, 19–28.

Rodrigues, D.M., Turchetto, C., Callegari-Jacques, S.M., and Freitas, L.B. (2018). Can the reproductive system of a rare and narrowly endemic plant species explain its high genetic diversity? Acta Botanica Brasilica *32*, 180–187.

Rodriguez-Saona, C., Parra, L., Quiroz, A., and Isaacs, R. (2011). Variation in highbush blueberry floral volatile profiles as a function of pollination status, cultivar, time of day and flower part: implications for flower visitation by bees. Ann Bot *107*, 1377–1390.

Sasse, J., Simon, S., Gübeli, C., Liu, G.-W., Cheng, X., Friml, J., Bouwmeester, H., Martinoia, E., and Borghi, L. (2015). Asymmetric localizations of the ABC transporter PaPDR1 trace paths of directional strigolactone transport. Curr. Biol. *25*, 647–655.

Sasse, J., Schlegel, M., Borghi, L., Ullrich, F., Lee, M., Liu, G.-W., Giner, J.-L., Kayser, O., Bigler, L., Martinoia, E., et al. (2016). *Petunia hybrida* PDR2 is involved in herbivore defense by controlling steroidal contents in trichomes. Plant Cell Environ. *39*, 2725–2739.

Schlüter, P.M., Xu, S., Gagliardini, V., Whittle, E., Shanklin, J., Grossniklaus, U., and Schiestl, F.P. (2011). Stearoyl-acyl carrier protein desaturases are associated with floral isolation in sexually deceptive orchids. PNAS *108*, 5696–5701.

Segatto, A.L.A., Ramos-Fregonezi, A.M.C., Bonatto, S.L., and Freitas, L.B. (2014). Molecular insights into the purple-flowered ancestor of garden petunias. Am. J. Bot. *101*, 119–127.

Sheehan, H., Moser, M., Klahre, U., Esfeld, K., Dell'Olivo, A., Mandel, T., Metzger, S., Vandenbussche, M., Freitas, L., and Kuhlemeier, C. (2016). *MYB-FL* controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nature Genetics *48*, 159–166.

Shuttleworth, A., and Johnson, S.D. (2010). The missing stink: sulphur compounds can mediate a shift between fly and wasp pollination systems. Proc Biol Sci *277*, 2811–2819.

Skaliter, O., Kitsberg, Y., Sharon, E., Shklarman, E., Shor, E., Masci, T., Yue, Y., Arien, Y., Tabach, Y., Shafir, S., et al. (2021). Spatial patterning of scent in petunia corolla is discriminated by bees and involves the ABCG1 transporter. The Plant Journal *106*, 1746–1758.

Sonnhammer, E., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. Proceedings. International Conference on Intelligent Systems for Molecular Biology *6*, 175—182.

Stehmann, J., and Semir, J. (2005). New species of *Calibrachoa* and *Petunia* (Solanaceae) from subtropical South America. Monographs in Systematic Botany from the Missouri Botanical Garden *104*, 341–348.

Stehmann, J.R., Lorenz-Lemke, A.P., Freitas, L.B., and Semir, J. (2009). The genus *Petunia*. In Petunia, (New York, NY: Springer New York), pp. 1–28.

Turchetto, C., Segatto, A.L.A., Silva-Arias, G.A., Beduschi, J., Kuhlemeier, C., Bonatto, S.L., and Freitas, L.B. (2019). Contact zones and their consequences: hybridization between two ecologically isolated wild *Petunia* species. Bot J Linn Soc *190*, 421–435.

Ueshima, R., and Asami, T. (2003). Single-gene speciation by left–right reversal. Nature *425*, 679–679.

Van Moerkercke, A., Galván-Ampudia, C.S., Verdonk, J.C., Haring, M.A., and Schuurink, R.C. (2012). Regulators of floral fragrance production and their target genes in petunia are not exclusively active in the epidermal cells of petals. J Exp Bot *63*, 3157–3171.

Verdonk, J.C., Ric de Vos, C.H., Verhoeven, H.A., Haring, M.A., van Tunen, A.J., and Schuurink, R.C. (2003). Regulation of floral scent production in petunia revealed by targeted metabolomics. Phytochemistry *62*, 997–1008.

Verdonk, J.C., Haring, M.A., van Tunen, A.J., and Schuurink, R.C. (2005). *ODORANT1* regulates fragrance biosynthesis in *Petunia* flowers. Plant Cell *17*, 1612–1624.

Vereecken, N.J., Cozzolino, S., and Schiestl, F.P. (2010). Hybrid floral scent novelty drives pollinator shift in sexually deceptive orchids. BMC Evolutionary Biology *10*, 103.

Waelti, M.O., Muhlemann, J.K., Widmer, A., and Schiestl, F.P. (2008). Floral odour and reproductive isolation in two species of *Silene*. J Evol Biol *21*, 111–121.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis (Springer-Verlag New York).

Widhalm, J.R., Jaini, R., Morgan, J.A., and Dudareva, N. (2015). Rethinking how volatiles are released from plant cells. Trends in Plant Science *20*, 545–550.

Xiong, J., Feng, J., Yuan, D., Zhou, J., and Miao, W. (2015). Tracing the structural evolution of eukaryotic ATP binding cassette transporter superfamily. Scientific Reports *5*, 16724.

Xu, S., Schlüter, P.M., Scopece, G., Breitkopf, H., Gross, K., Cozzolino, S., and Schiestl, F.P. (2011). Floral isolation is the main reproductive barrier among closely related sexually deceptive orchids. Evolution *65*, 2606–2620.

Yarahmadov, T., Robinson, S., Hanemian, M., Pulver, V., and Kuhlemeier, C. (2020). Identification of transcription factors controlling floral morphology in wild *Petunia* species with contrasting pollination syndromes. The Plant Journal *104*, 289–301.

Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. Current Protocols in Bioinformatics *69*, e96.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution *8*, 28–36.

Chapter 4
General discussion

The work presented in this manuscript focuses on pollination syndromes in the genus *Petunia*. In this genus, pollinators are important actors contributing to reproductive isolation between species (Dell'Olivo et al., 2011); the study of the genetic basis of phenotypic changes that affect pollinator-related traits can thus offer valuable insight into the genetic processes that have underlain the speciation of this genus (Dell'Olivo and Kuhlemeier, 2013).

The manuscript includes an introductory chapter (Chapter 1), two experimental chapters (Chapter 2 and 3) and a general discussion (Chapter 4). Chapter 1 introduces the topics and presents the *Petunia* genus. Chapter 2 focuses on the genetic architecture of pollinator-specific traits in *P. axillaris* and *P. exserta*, and takes advantage of secondary contact zones to investigate this topic. Chapter 3 is centred on the conservation of the *P. hybrida* scent transporter *ABCG1* in the wild petunia species, and on how the transporter could be associated to the changes in scent profiles that characterised the pollinator shifts. In the present chapter I discuss the results obtained, their limitations and their implications.

4.1 Pollination syndrome traits in secondary contact zones

When pollinators constitute a barrier to interspecific reproduction, the study of the phenotypic traits that define the pollination syndromes can help in the understanding of the evolutionary processes. In particular, studying the genetic basis of the pollination syndrome traits can inform us on the genetic targets of speciation. In the pollinator shift between moth-pollinated *P. axillaris* and hummingbird-pollinated *P. exserta* the genetic bases of the trait differences have been described, and display different genetic architectures.

4.1.1 Control of floral colour and the speciation island

Flower colour is generally considered a genetically simple trait, and changes in this trait have been often described as due to one or very few genes (Esfeld et al., 2018; Hoballah et al., 2007; Quattrocchio et al., 1999; Wessinger and Rausher, 2015; Yuan et al., 2013). Therefore, it is of particular interest that while the shift in UV absorbance (determined by quantity of flavonols) in *P. exserta* is underlain by a simple 1 bp deletion causing a shift in the reading frame of the transcription factor *MYB-FL*, the shift in visible colour involves multiple genetic modifications (Berardi et al., 2021; Sheehan et al., 2016). These findings show that traits traditionally considered simple can have a more complex genetic basis. The results of the genetic architecture analysis presented in Chapter 2 are in agreement with the previously described control of UV colour in these *Petunia* species but highlight a particular pattern in the variants associated to the phenotype in the GWAS: flavonol content is strongly associated to more than half the length of chromosome 2. The prediction of a simple genetic basis for flavonol content therefore seems to contradict the association study results.

A study based on artificial crosses between *P. axillaris* and *P. exserta* suggests that chromosome 2 hosts a multifunctional speciation island, i.e. a region of reduced recombination where loci responsible for reproductive isolation are grouped and are inherited together (Hermann et al., 2013). This region has been associated to the differences in scent, UV and visible colour, and morphology of reproductive

organs. The recombination rate in this region is low but not zero, and rare recombinants can be found that disrupt the inheritance of the group of traits. The authors therefore suggest that the loci may be experiencing suppression of recombination associated to heterochromatic regions. The broad region associated to flavonol content presented in Chapter 2 could therefore be motivated by the reduced recombination in the chromosome. Indeed, reduced recombination in pericentromeric and neighbouring regions has been observed to extend over more than half a chromosome length in several cases (for example de Haas et al., 2017; Rodgers-Melnick et al., 2015). On the other hand, the association of multiple phenotypic traits in the same chromosomal region reported by Hermann and colleagues (Hermann et al., 2013) seems to be absent in the contact zones analysed in this manuscript (except for scent production, which was not analysed in the contact zones). None of the phenotypes considered correlate with flavonol content, nor present a similar association pattern. I thus suggest that the cluster of pollination syndrome genes observed in the artificial crosses could have been disrupted by recombination during secondary contact.

4.1.2 Control of morphology: genetic complexity may not be a disadvantage

Continuous traits are generally considered to have a complex genetic basis. Supporting this classical idea in plants, the control of morphology and organ size is dependent on large and complex pathways and on interactions with phytohormones (Czesnick and Lenhard, 2015). Thus, the modification of morphological traits during evolution is often considered to require several changes of minor effect at the genetic level. Morphological speciation traits have been associated to polygenic or oligogenic bases, often involving loci of small to medium effect such as in the mating system change of *Leptosiphon* and *Mimulus*, and in the floral shape divergence in *Ipomopsis* (Fishman et al., 2002; Goodwillie et al., 2006; Nakazato et al., 2013). On the other hand, examples of morphological traits controlled by one gene only have been described too. In tomato, the evolution of self-pollination by means of a reduction in style length is due to a change in the ex-

pression of the transcription factor *LO2* (Chen et al., 2007). In *Aquilegia*, the presence of a nectar spur is dependent on a single transcription factor *POPOVICH* (Ballerini et al., 2020).

My study of floral morphology in *Petunia axillaris* and *P. exserta* indicates that pistil and tube length are controlled by several genes of medium to large effect, thus indicating an oligogenic basis.

Pistil length was previously associated to the speciation island on chromosome 2, and the distal section of the corolla tube (D2) had also been associated to this chromosome in a QTL study (Hermann et al., 2013, 2015). Thus, the region on chromosome 2 associated to the traits in my GWAS analysis find support in previous work but require additional investigation to verify coordinate overlap with the QTL, and to narrow down the associated region to the causative genes.

More recently, changes in pistil length in these species were shown to be influenced by the transcription factors *EOBI* and *EOBII*, situated on chromosome 4 and at the beginning of chromosome 2 respectively (Yarahmadov et al., 2020). Even though the region associated to the phenotype does not include these two genes, I cannot exclude a role for these two transcription factors in the control of the trait difference between the species. In fact, *EOBI* and *EOBII* are involved in several floral pathways such as anther dehiscence and male fertility, flavonol levels, scent production and flower opening in *Arabidopsis*, *Nicotiana* and *Petunia* (Colquhoun et al., 2011; Mandaokar et al., 2006; Spitzer-Rimon et al., 2010, 2012; Zhang et al., 2021). Hence, the modification of the expression of these transcription factors should be finely regulated to avoid developmental defects and other undesired effects. The results obtained in my work indicate that a region at the end of chromosome 2 is associated to this phenotype. I can thus hypothesise that a regulator of *EOBI* or *EOBII* or both is included in this region, and that a difference in the promoter region of these two transcription factors results in their differential expression in the two species. This hypothesis allows for a fine regulation of time and place of the expression by the action of the higher-level regulator.

Leaving hypotheses to the side, I can nonetheless conclude that the classical idea that morphological traits are polygenic traits is not always valid. Changes in morphology can be associated to few genes or many genes. In fact, the diversity in their genetic architectures may provide an advantage in the evolutionary process. This idea finds support in the recent work of Besnard and colleagues on *Caenorhabditis* (Besnard et al., 2020). *Caenorhabditis* evolves quickly and has a short life cycle thus allowing the observation of evolution "in real time". One cell in particular shows an extremely fast evolutionary rate in the reproductive system of this organism. The authors identified the causal mutations of the trait changes, and found that rather than being localised in specific regions of the genome, the causal mutations are scattered and affect distinct biological pathways. The authors thus come to the conclusion that the broad genetic basis that underlies cell fate in this organism is responsible for its fast evolutionary rate (Besnard et al., 2020).

The role of complex morphological or quantitative traits in the speciation process is therefore not to be underestimated, as these traits have the potential for quick divergence during speciation.

4.1.3 Selection in sympatric conditions may be targeting different traits than during divergence

In this study I analysed individuals of the species *P. axillaris* and *P. exserta* living in sympatry at two contact zones. The populations that are found in these areas include plants representative of the pure species and individuals genetically admixed at various levels (Caballero-Villalobos et al., 2021). The reproductive isolation between the species is therefore not complete and gene flow is possible, although species integrity seems to be maintained (i.e. gene flow has not homogenised the genetic components of all individuals in the regions, Caballero-Villalobos et al., 2021). Caballero-Villalobos and colleagues suggested that a combination of reproductive barriers is responsible for the maintenance of the species and may include abiotic factors.

My results show that two wide regions on chromosome 2 and 5 are under positive selection at the contact zones. Unfortunately, I am not able to distinguish if the selection pressure we detect is happening as a consequence of the secondary contact, or if it is due to the primary species divergence (Garner et al., 2018). This limitation could be overcome by the comparison of the selection scan presented in this manuscript with one obtained from allopatric populations of *P. axillaris* and *P. exserta*. I could thus discriminate between the action of divergence and that of secondary contact. In the case of a signal due to primary divergence during speciation, I could hypothesise that the regions are associated to traits that characterise the divergence between the species, such as adaptation to the environment or further specialisation in the pollination syndromes. In the case of selection only happening in secondary contact, I can instead hypothesise that the regions are associated to hybrid incompatibilities or to traits that reinforce the isolation of the species. The confirmation of these hypotheses requires further work, nonetheless my results indicate that these regions are under selection.

The selection signal on chromosome 2 overlaps with the region associated to pistil and tube length in the GWAS. The overlap of these two regions can lead us to further speculate that these two traits may be under selection. Several examples of selection on reproductive organs morphology have already been described. In *Ipomopsis*, Wolf and colleagues have advanced the hypothesis that efficiency of pollen transfer by hummingbirds between two species of this genus may be connected with the position and exertion of the reproductive organs (Wolf et al., 2001). Mechanical isolation between sympatric species is an important barrier in several genera (Kay, 2006; Kephart and Theiss, 2004), and particularly in *Achimenes* a role for floral morphology was suggested to provide reinforcement, here intended as the selection acting on traits to increase reproductive isolation in sympatry (Ramírez-Aguirre et al., 2019). In *Silene*, a multiyear study found that exertion of the stigma was under selection in a hummingbird-pollinated plant (Reynolds et al., 2010).

I can thus hypothesise that in *P. axillaris* and *P. exserta*, pistil and tube length are under selection as part of the species divergence process, or due to the renovated

need for reproductive isolation during secondary contact. A connection with phenotypic traits conferring reproductive isolation is not the only explanation possible for these selection signals. Hybrid incompatibilities may have developed during divergence and thus be selected upon in the secondary contact zones. Although plants from the two species are generally compatible, their hybrids may not be fit in the environment when a particular combination of loci is present (such as in Dobzhansky–Muller incompatibilities, Lindtke and Buerkle, 2015). This hypothesis should be carefully considered, in particular with attention to the effect of abiotic factors on the individuals.

Prominently absent in the regions under selection presented here is the locus of *MYB-FL*. *MYB-FL* is responsible for the difference in flavonol content in the flowers of *P. axillaris* and *P. exserta*, which results in UV absorbent flowers in the hawkmoth-pollinated species and in UV-reflective flowers in the hummingbird-pollinated species (Sheehan et al., 2016). A change in this trait alone has proven to affect the hawkmoth's preference and *MYB-FL* is therefore considered as a speciation gene (Hoballah et al., 2007; Sheehan et al., 2016). In the acquisition of red colour by *P. exserta,* the loss of MYB-FL activity is necessary to make the substrate available for the biosynthesis of anthocyanins (Berardi et al., 2021). The gene is thus important for both species. The absence of this locus from the regions under selection may be motivated by the loss of importance of the trait UV absorbance in maintaining reproductive isolation at the contact zones, in favour of a more relevant role for other floral traits or genetic incompatibilities.

My results show that a region associated to an important pollination syndrome trait (UV absorbance) is not under selection in the sympatric zones. Other regions of the genome are instead showing a signal of positive selection. While the results have some limitations in defining the timing of the detected selection, I suggest ways to improve the analysis and resolve current uncertainties. I also provide hypothesis on the causes that underlie these signals.

4.1.4 Perspectives

In discussing the results of this manuscript I have hinted at additional experiments that could help clarify the subject. A comparison of the reproductive isolation barriers between the allopatric and between the sympatric species could help define the role of pollinators in the evolutionary process. In particular, the precise quantification of each pre- and post-pollination barrier could inform on the selection acting is these regions. Pollinator essays could be used to measure the efficiency of pollen transfer between the species by the same pollinator, although the experiment would have to be carefully designed in order to account for the innate preference of the pollinator for one species. The isolation of the loci responsible for organ lengths would provide additional material to test pollinator efficiency differences. To isolate these loci we could take advantage of the work of Yarahmadov and colleagues (2020), who characterised pistil growth and associated it to transcriptome analysis. Flower development is a rather well studied topic, thus bibliographic resources on genes involved in these processes could be integrated. The involvement of candidate loci could be tested with CRISPR-Cas9 or introgression lines. Once identified the genes or loci, their evolutionary history and selection signatures could be analysed using different tools and resources such as sequence alignments and transcriptomes of other *Petunia* species, as well as data from wild individuals.

4.2 Scent transport goes wild

Scent is an important signal that attracts pollinators to the flowers, and it can affect pollinator choice even when other cues are present (Klahre et al., 2011; Raguso and Willis, 2002, 2005). In the *Petunia* genus, the evolution of different pollination syndromes was accompanied by changes in scent production (Stehmann et al., 2009). These changes are underlain by simple genetic modifications of a few enzymes and transcription factors that affect the biosynthesis of the volatiles. These changes in scent production result in differences in the emitted volatiles, suggesting that the route of scent emission could adapt to the production changes.

In the flower, volatiles can be emitted via passive diffusion through the plasma membrane, the cell wall and the cuticle, or via active transport (Dudareva and Pichersky, 2006). In *P. hybrida* cv Mitchell, scent emission depends on cuticle thickness and on the transporter *ABCG1* (Adebesin et al., 2017; Liao et al., 2020). The cuticle layer acts as a sink for the volatiles and a reduction in its thickness results in a reduction of the emitted volatiles. The action of the cuticle layer affects only compounds with low volatility, while the main scent compounds that characterise the wild *Petunia* species (methylbenzoate and benzaldehyde) are not affected (Liao et al., 2020).

The emission of these compounds is instead dependent on the presence of the plasma-membrane transporter ABCG1 in the same *P. hybrida* cv Mitchell (Adebesin et al., 2017). This research cultivar is the result of a cross between *P. axillaris* and the commercial cultivar *P. hybrida* Rose du Ciel, and its transcriptome includes a majority of sequences derived from *P. axillaris*, and a smaller proportion derived from a species of the short-tube clade such as *P. inflata* (Bombarely et al., 2016; Mitchell et al., 1980). The similarity in the volatiles emitted by *P. hybrida* cv Mitchell and *P. axillaris*, and the strong presence of transcripts of *P. axillaris*-origin in the research cultivar motivated us to hypothesise that the ABCG1 transporter could have a role in the wild petunias.

I found that the sequence of *P. hybrida* cv Mitchell ABCG1 is closer to *P. inflata*, suggesting that the gene encoding for it was inherited from the short-tube clade. In the commercial cultivars of *Petunia*, such as Rose du Ciel, the use of *P. axillaris* in crosses was motivated by the white colour of the flower which allowed the introgression of appealing colours. The inheritance of the transporter from a coloured *Petunia* could be hence due to hitchhiking of the gene with traits desirable for commercial uses.

While the sequence of the protein is rather conserved in the motif regions, the most striking result is that the gene is not expressed in the scented *P. axillaris*. *P. axillaris* uses scent to attract hawkmoths, and its main volatiles are methylbenzoate and benzaldehyde, two volatiles affected by the reduction in *ABCG1* expression in *P. hybrida* (Adebesin et al., 2017; Hoballah et al., 2005; Klahre et al., 2011).

These observations lead me to two alternative hypotheses: 1) *P. axillaris* uses another transporter to effect volatile emission, 2) *P. axillaris* does not need a transporter to emit these volatiles but relies solely on passive diffusion.

The first hypothesis is considered in Chapter 3. The characterisation of all members of the ABCG subfamily and the analysis of their expression identifies genes that have an expression pattern comparable to that of *ABCG1* in *P. hybrida*. I identify four genes of particular interest. These genes can be further investigated to validate their potential role in scent emission by means of CRISPR-Cas9 transformation. We cannot exclude that a transporter from a different family or subfamily could be involved though. To address this possibility we could extend the analysis of the expression pattern to the whole transcriptome of *P. axillaris*, although a careful design of the candidate selection has to be made in order to reduce sensibly the number of false positives. For example, we could integrate information on the predicted localisation of the protein sequences to select only those with plasma-membrane locality. The identification of an alternative gene responsible for scent transport through the plasma membrane would indicate that scent

emission is effected similarly in *P. hybrida* and *P. axillaris*. The use of a different gene could indicate redundancy in the role.

The second hypothesis instead addresses the option that active transport is not used in the scent emission of *P. axillaris*. In the experiments of Adebesin and colleagues, the reduction of *ABCG1* expression to 20% of its wild type level corresponds to a decrease of about 50% of the emission of methylbenzoate and benzaldehyde (Adebesin et al., 2017). The emission of methylbenzoate is particularly high in *P. hybrida* cv Mitchell (~ 150 nmol g_{FW}^{-1} h^{-1}), thus a reduction to 50% in methylbenzoate emission means that a considerable amount of this volatile is still emitted in the plants with a reduced expression of *ABCG1*. The direct comparison of the volatile emission between *P. axillaris* and *P. hybrida* cv Mitchell could help us in this direction. Skaliter and colleauges, in a study on the spatial distribution of scent on the floral corolla, reported the total scent emission of these two species and showed that *P. axillaris* emits less scent per tissue area than *P. hybrida* (Skaliter et al., 2021). This observation suggests the possibility that the quantity of volatiles produced by *P. axillaris* may in fact not need an active transporter to successfully reach the environment. To investigate this option in further detail we could compare directly the scent quantity and quality of *P. axillaris* and *P. hybrida* cv Mitchell. Once confirmed that *P. axillaris* emits less volatiles, the proof that this species needs no transporter is rather hard to achieve. One indirect proof could be provided by the observation of *ABCG1* expression in other cultivars of *P. hybrida* that display different levels of scent emission. The consistent correlation of high emission with high expression of the transporter and of low emission with low expression would hint in this direction.

While certain processes and pathways remained extremely conserved through speciation and commercial selection in *Petunia*, scent emission did not (Esfeld et al., 2018; Quattrocchio et al., 1999). The reason behind this observation could be a redundancy in the proteins that perform the task, or an essential difference in the emission process.

4.3 Concluding remarks

Understanding how evolution works poses several challenges. With this manuscript I hope to contribute to the knowledge of certain aspects of the process; in particular I suggest that morphological traits have the same potential as simpler traits for being controlled by speciation genes and thus quickly diverge. I hope that my results provide useful knowledge to the *Petunia* community and beyond, to all the people passionate about evolution.

4.4 References

Adebesin, F., Widhalm, J.R., Boachon, B., Lefèvre, F., Pierman, B., Lynch, J.H., Alam, I., Junqueira, B., Benke, R., Ray, S., et al. (2017). Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter. Science *356*, 1386–1388.

Ballerini, E.S., Min, Y., Edwards, M.B., Kramer, E.M., and Hodges, S.A. (2020). *POPOVICH*, encoding a C2H2 zinc-finger transcription factor, plays a central role in the development of a key innovation, floral nectar spurs, in *Aquilegia*. PNAS *117*, 22552–22560.

Berardi, A.E., Esfeld, K., Jäggi, L., Mandel, T., Cannarozzi, G.M., and Kuhlemeier, C. (2021). Complex evolution of novel red floral color in *Petunia*. Plant Cell *33*, 2273–2295.

Besnard, F., Picao-Osorio, J., Dubois, C., and Félix, M.-A. (2020). A broad mutational target explains a fast rate of phenotypic evolution. ELife *9*, e54928.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C.S., Bliek, M., Boersma, M.R., Borghi, L., Bruggmann, R., et al. (2016). Insight into the evolution of the Solanaceae from the parental genomes of *Petunia hybrida*. Nature Plants *2*, 16074.

Caballero-Villalobos, L., Silva-Arias, G.A., Turchetto, C., Giudicelli, G.C., Petzold, E., Bombarely, A., and Freitas, L.B. (2021). Neutral and adaptive genomic variation in hybrid zones of two ecologically diverged *Petunia* species (Solanaceae). Botanical Journal of the Linnean Society *196*, 100–122.

Chen, K.-Y., Cong, B., Wing, R., Vrebalov, J., and Tanksley, S.D. (2007). Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. Science *318*, 643–645.

Colquhoun, T.A., Schwieterman, M.L., Wedde, A.E., Schimmel, B.C.J., Marciniak, D.M., Verdonk, J.C., Kim, J.Y., Oh, Y., Gális, I., Baldwin, I.T., et al. (2011). *EOBII* controls flower opening by functioning as a general transcriptomic switch. Plant Physiology *156*, 974–984.

Czesnick, H., and Lenhard, M. (2015). Size control in plants—lessons from leaves and flowers. Cold Spring Harbor Perspectives in Biology *7*, a019190.

Dell'Olivo, A., and Kuhlemeier, C. (2013). Asymmetric effects of loss and gain of a floral trait on pollinator preference. Evolution *67*, 3023–3031.

Dell'Olivo, A., Hoballah, M.E., Gübitz, T., and Kuhlemeier, C. (2011). Isolation barriers between *Petunia axillaris* and *Petunia integrifolia* (Solanaceae). Evolution *65*, 1979–1991.

Dudareva, N., and Pichersky, E. (2006). Biology of floral scent (CRC Press).

Esfeld, K., Berardi, A.E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. (2018). Pseudogenization and resurrection of a speciation gene. Current Biology *28*, 3776-3786.e7.

Fishman, L., Kelly, A.J., and Willis, J.H. (2002). Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. Evolution *56*, 2138–2155.

Garner, A.G., Goulet, B.E., Farnitano, M.C., Molina-Henao, Y.F., and Hopkins, R. (2018). Genomic signatures of reinforcement. Genes *9*, 191.

Goodwillie, C., Ritland, C., and Ritland, K. (2006). The genetic basis of floral traits associated with mating system evolution in *Leptosiphon* (polemoniaceae): an analysis of quantitative trait loci. Evolution *60*, 491–504.

de Haas, L.S., Koopmans, R., Lelivelt, C.L.C., Ursem, R., Dirks, R., and Velikkakam James, G. (2017). Low-coverage resequencing detects meiotic recombination pattern and features in tomato RILs. DNA Research *24*, 549–558.

Hermann, K., Klahre, U., Moser, M., Sheehan, H., Mandel, T., and Kuhlemeier, C. (2013). Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. Current Biology *23*, 873–877.

Hermann, K., Klahre, U., Venail, J., Brandenburg, A., and Kuhlemeier, C. (2015). The genetics of reproductive organ morphology in two *Petunia* species with contrasting pollination syndromes. Planta *241*, 1241–1254.

Hoballah, M.E., Stuurman, J., Turlings, T.C.J., Guerin, P.M., Connétable, S., and Kuhlemeier, C. (2005). The composition and timing of flower odour emission by wild *Petunia axillaris* coincide with the antennal perception and nocturnal activity of the pollinator *Manduca sexta*. Planta *222*, 141–150.

Hoballah, M.E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., and Kuhlemeier, C. (2007). Single gene–mediated shift in pollinator attraction in *Petunia*. The Plant Cell *19*, 779–790.

Kay, K.M. (2006). Reproductive isolation between two closely related hummingbird pollinated neotropical gingers. Evolution *60*, 538–552.

Kephart, S., and Theiss, K. (2004). Pollinator-mediated isolation in sympatric milkweeds (*Asclepias*): do floral morphology and insect behavior influence species boundaries? New Phytologist *161*, 265–277.

Klahre, U., Gurba, A., Hermann, K., Saxenhofer, M., Bossolini, E., Guerin, P.M., and Kuhlemeier, C. (2011). Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. Current Biology *21*, 730–739.

Liao, P., Ray, S., Boachon, B., Lynch, J.H., Deshpande, A., McAdam, S., Morgan, J.A., and Dudareva, N. (2020). Cuticle thickness affects dynamics of volatile emission from petunia flowers. Nature Chemical Biology 1–8.

Lindtke, D., and Buerkle, C.A. (2015). The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution *69*, 1987–2004.

Mandaokar, A., Thines, B., Shin, B., Markus Lange, B., Choi, G., Koo, Y.J., Yoo, Y.J., Choi, Y.D., Choi, G., and Browse, J. (2006). Transcriptional regulators of stamen development in *Arabidopsis* identified by transcriptional profiling. The Plant Journal *46*, 984–1008.

Mitchell, A.Z., Hanson, M.R., Skvirsky, R.C., and Ausubel, F.M. (1980). Anther culture of *Petunia*: genotypes with high frequency of callus, root, or plantlet formation. Zeitschrift Für Pflanzenphysiologie *100*, 131–145.

Nakazato, T., Rieseberg, L.H., and Wood, T.E. (2013). The genetic basis of speciation in the *Giliopsis* lineage of *Ipomopsis* (Polemoniaceae). Heredity *111*, 227–237.

Quattrocchio, F., Wing, J., Woude, K. van der, Souer, E., Vetten, N. de, Mol, J., and Koes, R. (1999). Molecular analysis of the *anthocyanin2* gene of *Petunia* and its role in the evolution of flower color. The Plant Cell *11*, 1433–1444.

Raguso, R.A., and Willis, M.A. (2002). Synergy between visual and olfactory cues in nectar feeding by naïve hawkmoths, *Manduca sexta*. Animal Behaviour *64*, 685–695.

Raguso, R.A., and Willis, M.A. (2005). Synergy between visual and olfactory cues in nectar feeding by wild hawkmoths, *Manduca sexta*. Animal Behaviour *69*, 407–418.

Ramírez-Aguirre, E., Martén-Rodríguez, S., Quesada-Avila, G., Quesada, M., Martínez-Díaz, Y., Oyama, K., and Espinosa-García, F.J. (2019). Reproductive isolation among three sympatric *Achimenes* species: pre- and post-pollination components. American Journal of Botany *106*, 1021–1031.

Reynolds, R.J., Dudash, M.R., and Fenster, C.B. (2010). Multiyear study of multivariate linear and nonlinear phenotypic selection on floral traits of hummingbird-pollinated *Silene virginica*. Evolution *64*, 358–369.

Rodgers-Melnick, E., Bradbury, P.J., Elshire, R.J., Glaubitz, J.C., Acharya, C.B., Mitchell, S.E., Li, C., Li, Y., and Buckler, E.S. (2015). Recombination in diverse maize is stable, predictable, and associated with genetic load. PNAS *112*, 3823–3828.

Sheehan, H., Moser, M., Klahre, U., Esfeld, K., Dell'Olivo, A., Mandel, T., Metzger, S., Vandenbussche, M., Freitas, L., and Kuhlemeier, C. (2016). *MYB-FL* controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nature Genetics *48*, 159–166.

Skaliter, O., Kitsberg, Y., Sharon, E., Shklarman, E., Shor, E., Masci, T., Yue, Y., Arien, Y., Tabach, Y., Shafir, S., et al. (2021). Spatial patterning of scent in petunia corolla is discriminated by bees and involves the ABCG1 transporter. The Plant Journal *106*, 1746–1758.

Spitzer-Rimon, B., Marhevka, E., Barkai, O., Marton, I., Edelbaum, O., Masci, T., Prathapani, N.-K., Shklarman, E., Ovadis, M., and Vainstein, A. (2010). *EOBII*, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in *Petunia*. Plant Cell *22*, 1961–1976.

Spitzer-Rimon, B., Farhi, M., Albo, B., Cna'ani, A., Ben Zvi, M.M., Masci, T., Edelbaum, O., Yu, Y., Shklarman, E., Ovadis, M., et al. (2012). The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by activating *ODO1* and structural scent-related genes in *Petunia*. Plant Cell *24*, 5089–5105.

Stehmann, J.R., Lorenz-Lemke, A.P., Freitas, L.B., and Semir, J. (2009). The genus *Petunia*. In Petunia, (New York, NY: Springer New York), pp. 1–28.

Wessinger, C.A., and Rausher, M.D. (2015). Ecological transition predictably associated with gene degeneration. Molecular Biology and Evolution *32*, 347–354.

Wolf, P.G., Campbell, D.R., Waser, N.M., Sipes, S.D., Toler, T.R., and Archibald, J.K. (2001). Tests of pre- and postpollination barriers to hybridization between sympatric species of *Ipomopsis* (Polemoniaceae). American Journal of Botany *88*, 213–219.

Yarahmadov, T., Robinson, S., Hanemian, M., Pulver, V., and Kuhlemeier, C. (2020). Identification of transcription factors controlling floral morphology in wild *Petunia* species with contrasting pollination syndromes. The Plant Journal *104*, 289–301.

Yuan, Y.-W., Sagawa, J.M., Young, R.C., Christensen, B.J., and Bradshaw, H.D. (2013). Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics *194*, 255–263.

Zhang, X., He, Y., Li, L., Liu, H., and Hong, G. (2021). Involvement of the R2R3-MYB transcription factor MYB21 and its homologs in regulating flavonol accumulation in *Arabidopsis* stamen. Journal of Experimental Botany *72*, 4319–4332.

Declaration of consent

On the basis of Article 18 of the PromR Phil.-nat. 19

Name/First name:	Binaghi Marta
Registration number:	17-135-260
Study program:	PhD in Molecular Life Sciences
	Bachelor □ Master □ Dissertation ■
Title of the thesis:	Understanding genetic and phenotypic diversity in the speciation of *Petunia*: evolution of quantitative and qualitative traits
Supervisor:	Prof. Dr. Cris Kuhlemeier

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 literar of the University Act of September 5th, 1996 and Article 69 of the University Statute of June 7th, 2011 is authorized to revoke the doctoral degree awarded on the basis of this thesis. For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with theses submitted by others.

Bern, 23 November 2021