MYB-FL controls gain and loss of floral UV-absorbance, a key trait affecting pollinator preference and reproductive isolation

Hester Sheehan[1], Michel Moser[1], Ulrich Klahre[1], Korinna Esfeld[1], Alexandre Dell'Olivo[1], Therese Mandel[1], Sabine Metzger[2], Michiel Vandenbussche[3], Loreta Freitas[4] and Cris Kuhlemeier[1]

[1]Institute of Plant Sciences, University of Bern, Bern, Switzerland
[2]Cologne Biocenter, Cluster of Excellence on Plant Sciences, MS Platform, University of Cologne, Cologne, Germany
[3] UMR5667 (ENS de Lyon, CNRS, INRA, UCBL), Ecole Normale Supérieure de Lyon, Lyon, France
[4]Laboratory of Molecular Evolution, Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Corresponding author: cris.kuhlemeier@ips.unibe.ch

Abstract

Adaptations to new pollinators involve multiple floral traits, each requiring coordinated changes in multiple genes. Despite this genetic complexity, shifts in pollination syndromes have happened frequently during angiosperm evolution. Here we study the genetic basis of floral UV-absorbance, a key trait for attracting nocturnal pollinators. In *Petunia*, mutations in a single gene, *MYB-FL*, explain two transitions in UV-absorbance. A gain of UV-absorbance from bee- to moth-pollination was determined by a *cis*-regulatory mutation, whereas a frameshift mutation caused subsequent loss of UV-absorbance during the transition from moth- to hummingbird-pollination. The functional differences in *MYB-FL* provide insight into the process of speciation and clarify phylogenetic relationships between nascent species.

Animal-mediated pollination has been the hallmark of the rapid and successful radiation of the angiosperms[1,2]. Animal pollination is performed by specific guilds of pollinators, such as bees, butterflies, bats and birds which are associated with distinct sets of floral traits, often referred to as pollination syndromes[3]. Changes in pollinator groups can cause shifts in pollination syndromes, driving reproductive isolation and even speciation[4-7]. These shifts have happened frequently during the evolution of the angiosperms[2,3], for instance, the transition from bee to hummingbird pollination has happened at least ten times in the Solanaceae family alone[8]. Such transitions require coordinated alterations in multiple floral traits, such as color, scent, nectar and morphology. This raises questions about the molecular basis of these adaptive changes (for review, see [9]). Even assuming a simple genetic architecture, multiple genes must be involved in a shift between two pollination syndromes and they must change in a temporal sequence that avoids maladapted intermediate phenotypes. In some cases, single genes underlying individual traits have been identified and changes to these genes can already cause major phenotypic differences and alter pollinator behavior[10-15]. However, we are still far removed from a comprehensive molecular description of any pollination syndrome shift, or any other complex reproductive barriers.

The South American genus *Petunia* (Solanaceae) has undergone recent diversification[16] and comprises at least three distinct pollination syndromes[17,18]. The genus is separated into two major clades associated with the length of the floral tube[18]. The majority of species are members of the short-tubed clade and all have small, purple flowers that are pollinated by solitary bees[19]. In the long-tubed clade, different pollination syndromes are present. We focus on two of these species: *P. axillaris*, which displays a typical hawkmoth pollination syndrome with white flowers that produce abundant volatiles at dusk; and *P. exserta*, which has a characteristic hummingbird pollination syndrome exhibiting bright red, non-scented flowers with protruding reproductive organs[19]. All *Petunia* species can be intercrossed in the laboratory and produce fully fertile progeny indicating the absence of postzygotic barriers[20-23], therefore reproductive barriers must be predominantly prezygotic. As with many recently speciated genera, the direction of evolutionary transitions within *Petunia* is uncertain and recent phylogenetic studies only agree that the short-tubed clade is ancestral[18].

UV light is not perceived by humans, but is perceived by insects and other animal pollinators[24-28]. Insect-white flowers (those reflecting UV and visible wavelengths) are rare in nature and hawkmoth pollinators prefer UV-absorbing over UV-reflecting white flowers (reviewed in ref. 29)[30-33]. Moreover, pollinators can learn to associate differences in UV-absorption with differences in reward[34]. A major class of UV-absorbing compounds in plants are the flavonols, which absorb in the relevant range (near UV light) for animal pollinators, and also serve critical roles including defense from biotic and abiotic stresses[35]. Flavonols derive from phenylalanine, as do anthocyanins, a predominant visible pigments in plants, and they share the dihydroflavonols as common precursors[36]. *P. integrifolia* and *P. inflata,* representatives of the short-tubed clade, have UV-reflective corollas (this study)[19,20]. Within the long-tubed clade, almost all *P. axillaris* accessions have UV-absorbent corollas, whereas *P. exserta* is reflective[19,37]. Assuming that bee pollination represents the ancestral state in the genus[11,18], the earliest flowers must have been UV-reflective. One hypothesis is that moth pollination evolved directly from bee pollination. This would have involved a gain-of-function in UV-absorbance in *P. axillaris*, then a subsequent loss in *P. exserta*. An alternative hypothesis is that hummingbird pollination evolved from bee pollination, a common

transition[38,39]. This would have required no changes in UV-absorbance before the gain of UV absorbance during the emergence of *P. axillaris*.

Knowledge of the molecular basis of differences in UV-absorbance between *P. axillaris* and *P. exserta* will help us resolve their evolutionary trajectory. A good example of this kind of approach is the *Petunia AN2* gene, which controls anthocyanin pigmentation in the short-tubed species[40]. The coding sequence of *AN2* is mutated in *P. axillaris* causing the petals to be white and pollinator attraction to shift from bees towards hawkmoths[11,40,41]. Since multiple, independent, non-functional alleles have been found in natural populations, the most parsimonious assumption is that white flowers and thus, *P. axillaris*, are derived.

In this work, we demonstrate the importance of UV-absorbance for hawkmoth attraction in *Petunia* and elucidate the molecular basis of differences in this trait. We show that a single gene, *MYB-FL*, is the major determinant of flavonol levels and, thus, UV-absorption between *P. inflata*, *P. axillaris* and *P. exserta.* The gain of UV-absorbance from *P. inflata* to *P. axillaris* is caused by upregulation of the *MYB-FL* promoter, whereas a frameshift mutation in *MYB-FL* is responsible for the difference in UV-absorbance between *P. axillaris* and *P. exserta*.

Results

Hawkmoths prefer to feed from UV-absorbent flowers

In order to assess the importance of floral UV-absorption, we presented the natural hawkmoth pollinator, *Manduca sexta*[11,42], with *P. axillaris* accessions that differ in floral UV-absorbance but not visible absorbance (Fig. 1a,b and Supplementary Fig. 1a,b). Other floral traits important for pollinator attraction did not differ substantially between these lines (Supplementary Fig. 1c-e). *M. sexta* preferred the UV-absorbent accession on first choice and the number of feeding events per line was also higher (Fig. 1c,d). No differences were observed in the time spent feeding per feeding event (Fig. 1e), indicating that the lack of UV-absorbance affects flower visibility but not access to the nectar. This experiment demonstrates that hawkmoths recognize and prefer UV-absorbent *P. axillaris* flowers, and that UV-reflective *P. axillaris* accessions may experience lower visitation and reproductive fitness if they rely on hawkmoths for pollination.

Differences in flavonol levels localize to a major QTL

P. inflata, *P. axillaris* and *P. exserta* flowers differ in appearance in visible and UV light (Fig. 2a). This difference in UV light is determined by contrasting levels of flavonols in their corolla limbs (Fig. 2b). In interspecific F_1 plants, flavonol levels are as high, or nearly as high, as in *P. axillaris* indicating that UV-absorbance is a (semi)-dominant trait (Fig. 2a,b). Previous work using an interspecific F_2 cross between *P. integrifolia* and *P. axillaris* mapped flavonol levels to a locus on chromosome II[20]. We carried out a QTL analysis of an F_2 population between *P. axillaris* and *P. exserta* which similarly revealed a major locus on chromosome II that explained 79% of the variation in flavonol levels between the two species, and a minor locus on chromosome III that explained 14% of the variation (Fig. 2c; Supplementary Table 1). This

is in agreement with our previous experiments that implicated a locus on chromosome II between *P. axillaris* and *P. exserta* that segregates for both UV color and visible color[37]. In our F_2 population, we also saw a negative correlation between UV-absorbance and visible absorbance (Fig. 2d), suggesting a trade-off between flavonols and anthocyanins. In order to determine whether the same gene or different genes are involved, we crossed the two reflective accessions, *P. inflata* with *P. exserta*, and saw no increase in UV-absorbance in the F_1 (Fig. 2e). This lack of genetic complementation is direct evidence that the same gene is involved in determining a lower level of UV-absorbance, and thus flavonols, in the two species.

A regulator of *FLS* determines differences in flavonols

An immediate candidate for determining the differences in flavonol levels between *P. inflata*, *P. axillaris* and *P. exserta*, was *flavonol synthase* (*FLS*), as it encodes the enzyme responsible for converting dihydroflavonol precursors into flavonols and maps to chromosome II[20,43]. The *FLS* gene was differentially expressed (Fig. 3a-c) and is closely linked to the major QTL on chromosome II (Supplementary Table 1). The predicted FLS protein sequences are very similar and likely to be functional in all three species (Supplementary Fig. 2a). Expression differences may be attributed to polymorphisms in the promoter region of *FLS* (*cis*-regulatory differences) or in a transcription factor that binds to the *FLS* promoter (*trans*-regulatory differences). Using RNA-sequencing (RNA-seq) data from genetic material in which the *FLS* gene is in a heterozygous state, only minor differences in expression between the parental alleles were detected (Fig. 3d,e). Altogether, these results implicate a *trans*-regulator of *FLS*, closely linked to *FLS* on chromosome II, as determining differences in flavonol levels between *P. inflata* and *P. axillaris*, and also between *P. axillaris* and *P. exserta*. In order to identify this *trans*-regulator of *FLS*, we focused on the transition between *P. axillaris* and *P. exserta*, because the difference in *FLS* expression is more pronounced (56-fold cf. 8-fold; Fig. 3b cf. Fig. 3a).

Mutations in *MYB-FL* affect UV-absorbance and visible color

A transposon mutagenesis screen using *P. hybrida* W138 (W138), a line with active transposable elements, was undertaken to isolate the *trans*-regulator of *FLS*. W138 is UV-reflective and does not complement the UV-absorbance phenotype when crossed to *P. exserta* (Fig. 2e). Therefore we assumed that it does not have a functional allele of the *trans*-regulator of *FLS*. W138 was crossed with *P. axillaris* as well as with the introgression line, IL2-1^{Pax,} which is in the *P. exserta* background and homozygous for *P. axillaris* over the QTL region (see Online Methods and ref. 37).Two F_1 populations of approximately 200 plants each were produced. These populations consisted of plants with UV-absorbent, pale-purple flowers and were screened for UV-reflective somatic sectors (for details, see Online Methods). We obtained a class of sectors from both populations which reflected UV light greatly compared to the surrounding F_1 tissue (Fig. 4a) and that were bright pink in visible light (Fig. 4b; Supplementary Fig. 3a-d). These sectors showed a reduction in flavonols similar to the level shown by W138 (Fig. 4c).

Identifying the causal mutation requires finding a gene that is disrupted by a transposon in the sector but not in the surrounding corolla tissue. Using a combination of PCR screening and transposon display (see Online Methods), we detected five independent transposon insertions in a gene, *MYB-FL*, with 38.7% protein identity to *Arabidopsis thaliana* MYB111, and with an uncommon R2R3-MYB exon structure of

four exons and three introns (Fig. 4d, Supplementary Table 2; Supplementary Fig. 4). *AtMYB111* is one member of a subfamily of R2R3-MYB transcription factors that regulate *FLS* expression in *Arabidopsis*[44,45]. Genotyping of the *P. axillaris* x *P. exserta* F_2 population showed that *MYB-FL* is fully linked to *FLS* (Supplementary Table 3). Two homologues of MYB-FL with 30.4% and 35.5% protein identity respectively were found in the *P. axillaris* genome (and homologues of all three of these genes were found in the *P. inflata* genome). *MYB-FL* is the only corolla-expressed gene, therefore, we focused on *MYB-FL*.

To confirm that *MYB-FL* is responsible for the UV-absorbance difference with *P. exserta*, we generated a germline *MYB-FL* mutation from a large somatic sector by regenerating plantlets from sepal tissue. A germline transposon insertion line, A1-95, was obtained with the *P. axillaris* allele of *MYB-FL* disrupted by a *dTph4* transposon ($MYB\text{-}FL^{Pax\text{-}dTph4}$) and the other *MYB-FL* allele from W138 (Supplementary Fig. 3e). This plant was crossed with *P. exserta*. Resulting progeny with genotype $MYB\text{-}FL^{Pax\text{-}dTph4}/MYB\text{-}FL^{Pex}$ were UV- reflective indicating that no complementation occurred and that $MYB\text{-}FL^{Pax\text{-}dTph4}$ is allelic with the causal *P. exserta* gene (Fig. 4e). Occasional UV-absorbent sectors were due to transposon excision (Supplementary Fig. 3f,g).

We also measured absorbance in the visible range in the bright pink sectors as they clearly appeared darker and a different hue to the unaffected surrounding tissue (Fig. 4b). Visible absorbance was increased by 1.5- to 4-fold in the sectors compared to the surrounding tissue (Supplementary Fig. 3a-d). Measurement of sample A1-95 by mass spectrometry confirmed the increase in anthocyanidins as well as an alteration in proportions of these compounds (Fig. 4f). Thus, the negative correlation between flavonol and anthocyanin content in the F_2 mapping population (Fig. 2d) may be attributed to the allelic status of *MYB-FL*.

MYB-FL is expressed but mutated in _P. exserta_

P. axillaris and *P. exserta* buds showed peak expression of *MYB-FL* at the mid stages of bud development, approximately the same time or a bit earlier than *FLS* expression (Fig. 4g and Fig. 3c), consistent with a role in regulating *FLS* expression. *P. exserta MYB-FL* and *FLS* expression are consistently lower than in *P. axillaris*. This is supported by RNA-seq analysis (Table 1). However, differences in expression of *MYB-FL* are moderate and are unlikely to explain the extreme differences in *FLS* expression between the different species. Therefore, we asked whether a functional difference in MYB-FL between the two species could explain this difference in *FLS* expression. A comparison of the full-length sequences of *MYB-FL* showed that the *P. exserta* allele has a single base-pair deletion in the third exon causing a frameshift and a truncated protein sequence (Fig. 4h, Supplementary Fig. 4).

Considering the negative correlation between absorbance in the UV and visible ranges (Fig. 2d and Fig. 4a-c,f) it could be that *MYB-FL* has a dual function as an activator of flavonol synthesis and an inhibitor of the anthocyanin pathway. For instance, as well as activating *FLS* expression, *MYB-FL* might inhibit the subset of flavonoid biosynthetic genes upregulated during anthocyanin production (*CHS-J*, *DFR*, *ANS*, *RT*, *MT*, *GST)*[46]. We investigated whether the anthocyanin-specific genes show expression differences depending on the presence or absence of the *P. axillaris* allele of *MYB-FL* ($MYB\text{-}FL^{Pax}$), by comparing RNA-seq data between $IL2\text{-}1^{Pax}$ and *P. exserta* (Table 1). This showed that the portion of the genome containing *MYB-FL* caused minor (10-30%) downregulation of *DFR*, *RT*, *MT* and *GST*. This suggests that

MYB-FL^{Pax} plays at most a minor role in regulating the anthocyanin branch of the pathway between *P. axillaris* and *P. exserta* and does not explain the complete difference in anthocyanin levels between these species. Since *MYB-FL* appears to have minimal influence on anthocyanin biosynthetic genes, competition for the common substrates, the dihydroflavonols, is most plausible (e.g., refs. 43,47).

A *cis*-acting mutation upregulates *MYB-FL* in *P. axillaris*

The results presented above clearly show that *MYB-FL* determines a major difference in flavonol levels between *P. axillaris* and *P. exserta*. Complementation tests show that the same locus is also responsible for UV-reflectance in *P. inflata* (Fig. 2e). Thus we set out to investigate *MYB-FL* in *P. inflata*. *P. inflata MYB-FL* encodes a full-length protein sequence presumably producing a functional protein (Supplementary Fig. 4). *MYB-FL* expression is 24-fold lower in *P. inflata* than *P. axillaris* (Fig. 5a) which could explain the lower expression of *FLS* in *P. inflata* (Fig. 3a). To determine if *MYB-FL* showed *cis*- or *trans*-regulatory differences, we measured allele-specific expression in the *P. axillaris* x *P. inflata* F_1 and found that expression is 15-fold higher from the *P. axillaris* allele (Fig. 5b). This indicates a *cis*-regulatory difference in the regulation of *MYB-FL* between the two species.

We compared the upstream sequences of the *MYB-FL* alleles in *P. inflata*, *P. axillaris* and *P. exserta* as well as the intron regions and found three major structural polymorphisms (Fig. 5c). Most notably, a 977 bp insertion is present in the *P. axillaris* and *P. exserta* sequences, 12 bp upstream of the predicted TATA-box. Using bioinformatics prediction of *cis*-regulatory elements, we found a number of potential transcription factor binding sites inside this insertion region (Supplementary Table 4). Altogether, the genetic data, expression differences, and sequence differences show that *cis*-regulatory elements in the *MYB-FL* gene are likely to be responsible for the differences in expression of the two alleles and thus for differing flavonol levels in the petals of the two species. Future experiments such as promoter-reporter assays will narrow down the crucial regions that determine the expression differences.

***MYB-FL* differentiates species in natural populations**

P. exserta has a narrow distribution and is exclusively found in shallow shelters formed in sandstone towers. *P. axillaris* grows around but not inside these shelters[48,49]. Here we ask how the single base-pair deletion in the *MYB-FL P. exserta* reference accession is distributed in natural populations. *P. exserta* accessions were sampled from nine towers representing the distribution of the species in the Serra do Sudeste region of Brazil, whilst *P. axillaris* accessions were sampled from multiple locations in Argentina, Uruguay and Brazil, including regions of sympatry with *P. exserta* (Supplementary Table 5)[49] and genotyped using a CAPS marker. Out of 140 *P. exserta* accessions genotyped, 138 were homozygous for the single base-pair deletion. One individual could not be genotyped (see Supplementary Note) and one showed a homozygous *P. axillaris* genotype (*P. exserta* population 26; Supplementary Table 5).

Out of 159 *P. axillaris* accessions, 155 were homozygous for the *P. axillaris* allele and three individuals were heterozygous for the deletion. Two of these individuals came from a population that is sympatric with *P. exserta* (*P. axillaris* population 66; Supplementary Table 5). The other heterozygous individual came from a population in Uruguay (*P. axillaris* population 12; Supplementary Table 5), approximately 300 km removed from *P. exserta* populations. One individual could not be genotyped (see

Supplementary Note). Thus, the *MYB-FL* marker shows an almost perfect association between genotype and species classification.

We also tested individuals that came from a population, Pedra da Cruz, that included putative hybrids of *P. axillaris* and *P. exserta* (Supplementary Table 5). From these individuals, eight showed a homozygous *P. exserta* genotype, eight a homozygous *P. axillaris* genotype and five were heterozygous. Six individuals could not be accurately genotyped (see Supplementary Note). We did not have UV-absorbance information for any of these individuals, so we could not examine the association between this trait and *MYB-FL* genotype. Instead, we made use of seeds collected from plants from the same population as well as from a second population with putative hybrid individuals, Ponto 143 (Supplementary Table 5). Progenies were obtained with one to three of the possible genotypic classes (Fig. 6). *P. exserta* homozygous classes showed a low level UV-absorbance, and heterozygous or homozygous *P. axillaris* classes showed a comparatively higher level of UV-absorbance. This association was significant in the three families segregating for all three genotypic classes (Fig. 6).

Discussion

We have identified a single R2R3-type MYB transcription factor, MYB-FL, which, through different molecular mechanisms, is responsible for the high levels of UV-absorbing floral flavonols in *P. axillaris* and the low levels in *P. inflata* and *P. exserta*. With support from pollinator choice assays, we propose that high levels of flavonols evolved in *P. axillaris* to attract nocturnal visitors and were lost again during adaptation to hummingbird pollination. Changes in UV-absorbance are a key component of switches between pollination syndromes and, consequently, *MYB-FL* can be designated a speciation gene.

Flavonols were already present in early land plants, long before the evolution of flowers, and their ancestral function was likely in protection from light stress[50]. Likewise, R2R3-MYBs are present in early plants and presumably regulate flavonoid biosynthetic genes[51]. Duplications within this gene family are associated with new expression domains in higher plants[52-54]. Therefore, we assume that *MYB-FL* acquired its function in pollinator attraction through a similar process. *MYB-FL* joins other R2R3-MYBs involved in pollination syndrome shifts, such as transitions in anthocyanin levels and scent volatiles[11-13,40,52,53,55]. Our results further support the thought-provoking hypothesis that, like the MADS-box transcription factors provide the toolbox for floral whorl identity, R2R3-MYBs are preferentially recruited for shifts in floral pollination syndromes[56].

Phylogenetic relationships between nascent species are often difficult to resolve due to low genetic divergence, standing variation and continuing gene flow[57]. This is also true for *Petunia*, where the relationships in the long-tubed clade are especially uncertain[18,48,49]. To some extent the problem can be alleviated by high-throughput sequencing in combination with sophisticated statistical methods[58]. In complement to these approaches, we propose that examination of the functionally relevant polymorphisms – those that determine adaptive differences– can help resolve phylogenetic uncertainties. Based on our molecular analysis of *MYB-FL* function, the most parsimonious trajectory is that *MYB-FL* was originally a lowly expressed functional gene in the bee-pollinated ancestor, and then

acquired a *cis*-regulatory mutation that boosted its expression in *P. axillaris.* Subsequently, the gene became inactive in *P. exserta*, not through reduced expression, but due to a frameshift mutation.

The trade-off between anthocyanins and flavonols also suggests a most likely order of changes in these two traits. During the transition from *P. axillaris* to *P. exserta*, a decrease in flavonols must have occurred first, before the major increase in anthocyanin levels could have been achieved. Whether or not this increase would have occurred immediately would have depended on the expression level and functional capacity of the anthocyanin biosynthetic genes in this ancestral *P. axillaris* species. Rather than degeneration of a function no longer needed, loss of UV-absorbance may have been a prerequisite to produce the red color prevalent in hummingbird-pollinated flowers[59,60]. Furthermore, if the upregulation of *MYB-FL* and flavonols was the primary event during the transition from the short-tubed, bee-pollinated clade to *P. axillaris*, this could have caused loss of selection on maintenance of *AN2* function, something suggested by the occurrence of multiple nonsense alleles of this gene[11,40]. This highlights that genetic and biochemical mechanisms may constrain the types of transitions that can occur, and the sequence of trait changes within a single transition.

Genes involved in pre-zygotic reproductive isolation and speciation are being identified in a variety of plant and animal systems[14,61-64]. Pollination syndromes pose an interesting challenge. QTL analyses in *Petunia* suggest a total of at least a dozen loci explaining major differences in floral color, UV-absorbance, scent production, nectar production and morphology[13,65-67], comparable to other QTL studies[68-71]. Considering this genetic complexity, how can shifts in pollination syndromes have happened so frequently during angiosperm evolution? We suggest the following mechanisms. Firstly, evolutionary theory allows for initial large-effect mutations during adaptive bouts with fine-tuning through smaller effect changes[9,72]. Here, we show that differences in UV-absorption are specified by a single transcription factor, MYB-FL. Similar results were obtained for transitions in anthocyanin-based color[11,12,40,52,53,55]. If this trend holds, the number of genes will still be large but not near-infinite. Secondly, the *MYB-FL* gene is tightly linked to major loci for scent production and morphology[37]. Such an architecture will maintain favorable gene combinations in sympatry[73]. Thirdly, due to competition for a common substrate, upregulation of *MYB-FL* expression will enhance UV-absorption and make petals visibly paler. Thus, a single gene mutation can pleiotropically affect two traits. Finally, even changes in single genes can strongly affect pollinator preference (Fig. 1)[11]. In the case of *AN2*, loss of function increases attraction of hawkmoths without reducing visitation by the established bee pollinator[41]. Such intermediate stages with relaxed pollinator specificity, but enhanced reproductive assurance, need not be maladaptive during periods of rapid change.

Identifying the functionally relevant polymorphisms underlying all major QTL and determining their genomic architecture is now feasible. This will deepen our insight into the process of speciation and thereby resolve phylogenetic relationships between nascent species.

URLs

ImageJ software, http://www.imagej.org; Primer3, http://primer3.ut.ee/; Qgene QTL software, http://www.qgene.org/qgene/index.php; Kuhlemeier lab website,

http://www.ips.unibe.ch/deve/caps/index.html; R software, http://www.R-project.org/; FastqMcf, http://code.google.com/p/ea-utils/wiki/FastqMcf.

Accession codes

All RNA sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProjects PRJNA300613 and PRJNA300556. Gene sequences can be found in the GenBank/EMBL databases under the following accession numbers: *P. axillaris FLS*, KT962952; *P. exserta FLS*, KT962953; *P. inflata FLS*, KT962954; *P. axillaris MYB-FL*, KT962949; *P. exserta MYB-FL*, KT962950; *P. inflata MYB-FL*, KT962951; *P. axillaris MYB-FL* transposon-insertion alleles (see Supplementary Table 2), KT962941-KT962948.

Acknowledgments

We thank Christopher Ball, Jasmin Sekulovski, Nicole Signer and Ramon Zimmermann for expert care of our plants; Moritz Saxenhofer and Anna Feller for help with genotyping and screening transposon populations; Patrice Morel and Jan Zethof for help with the amplification and sequencing of *dTph1* transposon flanking sequences; Ana Luiza Cazé for help with phenotyping hybrid progeny plants; and Ana Lucia Segatto and Caroline Turchetto for field collections and DNA extraction. Kimberley Snowden, Avichai Amrad, Katrin Hermann, Holly Summers and Sarah Robinson provided constructive comments on manuscript drafts; Avichai Amrad, Ronald Koes, Jeroen Stuurman, and Tom Gerats provided advice and suggestions throughout the project; and Roman Köpfli assisted with figures. MV was financially supported by an ATIP-AVENIR (CNRS) and ANR-BLANC grant. LF was financed by Science without Borders, CNPq. The remaining authors were financed by the National Centre of Competence in Research "Plant Survival" and grants from the Swiss National Science Foundation and the University of Bern.

Author Contributions

Conceptualization, HS, UK, LF and CK; Methodology, HS, MM, UK, AD, MV, and CK; Software, MM; Formal analysis, HS, MM, and UK; Investigation, HS, MM, UK, AD, KE, TM, SM and MV; Resources, MV, LF and CK; Writing, HS, MM, SM, MV, and CK; Visualization, HS; Funding acquisition, CK.

References for main text

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

2. Faegri, K. & van der Pijl, L. *The Principles of Pollination Ecology*, (Pergamon Press, 1979).

3. Fenster, C.B., Armbruster, W.S., Wilson, P., Dudash, M.R. & Thomson, J.D. Pollination syndromes and floral specialization. *Annual Review of Ecology, Evolution, and Systematics* **35**, 375-403 (2004).

4. Kay, K. & Schemske, D. Pollinator assemblages and visitation rates for 11 species of neotropical *Costus* (Costaceae). *Biotropica* **35**, 198-207 (2003).

5. Gegear, R.J. & Burns, J.G. The birds, the bees, and the virtual flowers: can pollinator behavior drive ecological speciation in flowering plants? *The American Naturalist* **170**, 551-566 (2007).

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

7. Hopkins, R. & Rausher, M.D. Pollinator-mediated selection on flower color allele drives reinforcement. *Science* **335**, 1090-1092 (2012).

8. Knapp, S. On 'various contrivances': pollination, phylogeny and flower form in the Solanaceae. *Philosophical Transactions of the Royal Society B* **365**, 449-460 (2010).

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

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

11. Hoballah, M.E. *et al.* Single gene-mediated shift in pollinator attraction in *Petunia*. *The Plant Cell* **19**, 779-790 (2007).

12. Hopkins, R. & Rausher, M.D. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* **469**, 411-414 (2011).

13. Klahre, U. *et al.* Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Current Biology* **21**, 730-739 (2011).

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

15. Shang, Y. *et al.* The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. *New Phytologist* **189**, 602-615 (2011).

16. Lorenz-Lemke, A.P. *et al.* Diversification of plant species in a subtropical region of eastern South American highlands: a phylogeographic perspective on native *Petunia* (Solanaceae). *Molecular Ecology* **19**, 5240-5251 (2010).

17. Stehmann, J.R., Lorenz-Lemke, A.P., Freitas, L.B. & Semir, J. in *Petunia: Evolutionary, Developmental and Physiological Genetics* (eds. Gerats, T. & Strommer, J.) 1-28 (Springer, 2009).

18. Reck-Kortmann, M. *et al.* Multilocus phylogeny reconstruction: new insights into the evolutionary history of the genus *Petunia*. *Molecular Phylogenetics and Evolution* **81**, 19-28 (2014).

19. Gübitz, T., Hoballah, M.E., Dell'Olivo, A. & Kuhlemeier, C. in *Petunia: Evolutionary, Developmental and Physiological Genetics* (eds. Gerats, T. & Strommer, J.) 29-50 (Springer, 2009).

20. Wijsman, H.J.W. On the interrelationships of certain species of Petunia II. Experimental data: crosses between different taxa. *Acta Botanica Neerlandica* **32**, 97-107 (1983).

21. Ando, T. *et al.* Reproductive isolation in a native population of *Petunia sensu* Jussieu (Solanaceae). *Annals of Botany* **88**, 403-413 (2001).

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

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

24. Goldsmith, T.H. Hummingbirds see near ultra-violet light. *Science* **207**, 786-788 (1980).

25. White, R.H., Brown, P.K., Hurley, A.K. & Bennett, R.R. Rhodopsins, retinula cell ultrastructure, and receptor potentials in the developing pupal eye of the moth *Manduca sexta*. *Journal of Comparative Physiology A* **150**, 153-163 (1983).

26. Peitsch, D. *et al.* The spectral input systems of hymenopteran insects and their receptor-based colour vision. *Journal of Comparative Physiology A* **170**, 23-40 (1992).

27. Winter, Y., Lopez, J. & von Helversen, O. Ultraviolet vision in a bat. *Nature* **425**, 612-614 (2003).

28. Herrera, G. *et al.* Spectral sensitivities of photoreceptors and their role in colour discrimination in the green-backed firecrown hummingbird (*Sephanoides sephaniodes*). *Journal of Comparative Physiology A* **194**, 785-794 (2008).

29. Kevan, P., Giurfa, M. & Chittka, L. Why are there so many and so few white flowers? *Trends in Plant Science* **1**, 280-284 (1996).

30. Guldberg, L.D. & Atsatt, P.R. Frequency of reflection and absorption of ultraviolet light in flowering plants. *American Midland Naturalist* **93**, 35-43 (1975).

31. Chittka, L., Shmida, A., Troje, N. & Menzel, R. Ultraviolet as a component of flower reflections, and the colour perception of hymenoptera. *Vision Research* **34**, 1489-1508 (1994).

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

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

34. Kelber, A. & Hénique, U. Trichromatic colour vision in the hummingbird hawkmoth, *Macroglossum stellatarum* L. *Journal of Comparative Physiology A* **184**, 535-541 (1999).

35. Gould, K.S. & Lister, C. in *Flavonoids: Chemistry, Biochemistry and Applications* (eds. Andersen, Ø.M. & Markham, K.R.) 397-441 (CRC Press, 2006).

36. Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* **126**, 485-493 (2001).

37. Hermann, K. *et al.* Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. *Current Biology* **23**, 873-877 (2013).

38. Rausher, M.D. Evolutionary transitions in floral color. *International Journal of Plant Science* **169**, 7-21 (2008).

39. Thomson, J.D. & Wilson, P. Explaining evolutionary shifts between bee and hummingbird pollination: convergence, divergence and directionality. *International Journal of Plant Science* **169**, 23-38 (2008).

40. Quattrocchio, F. *et al.* Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. *The Plant Cell* **11**, 1433-1444 (1999).

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

42. Venail, J., Dell'Olivo, A. & Kuhlemeier, C. Speciation genes in the genus *Petunia*. *Philosophical Transactions of the Royal Society B* **365**, 461-468 (2010).

43. Holton, T.A., Brugliera, F. & Tanaka, Y. Cloning and expression of flavonol synthase from *Petunia hybrida*. *The Plant Journal* **4**, 1003-1010 (1993).

44. Mehrtens, F., Kranz, H., Bednarek, P. & Weisshaar, B. The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* **138**, 1083-1096 (2005).

45. Stracke, R. *et al.* Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal* **50**, 660-677 (2007).

46. Quattrocchio, F., Wing, J.F., Leppen, H.T.C., Mol, J.N.M. & Koes, R.E. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *The Plant Cell* **5**, 1497-1512 (1993).

47. Davies, K.M. *et al.* Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. *Euphytica* **131**, 259-268 (2003).

48. Lorenz-Lemke, A.P. *et al.* Diversity and natural hybridization in a highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. *Molecular Ecology* **15**, 4487-4497 (2006).

49. Segatto, A.L.A. *et al.* 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 (2014).

50. Pollastri, S. & Tattini, M. Flavonols: old compounds for old roles. *Annals of Botany* **108**, 1225-1233 (2011).

51. Rausher, M.D. in *The Science of Flavonoids* (ed. Grotewold, E.) 175-211 (Springer, 2006).

52. Cooley, A.M., Modliszewski, J.L., Rommel, M.L. & Willis, J.H. Gene duplication in *Mimulus* underlies parallel floral evolution via independent *trans*-regulatory changes. *Current Biology* **21**, 700-704 (2011).

53. Schwinn, K. *et al.* A small family of *MYB*-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *The Plant Cell* **18**, 831-851 (2006).

54. Albert, N.W. *et al.* A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *The Plant Cell* **26**, 962-980 (2014).

55. Streisfeld, M.A., Young, W.N. & Sobel, J.M. Divergent selection drives genetic differentiation in an R2R3-MYB transcription factor that contributes to incipient speciation in *Mimulus aurantiacus*. *PLOS Genetics* **9**, e1003385 (2013).

56. Yuan, Y., Byers, K.J.R.P. & Bradshaw Jr, H.D. The genetic control of flower–pollinator specificity. *Current Opinion in Plant Biology* **16**, 422-428 (2013).

57. Edwards, S.V. Is a new and general theory of molecular systematics emerging? *Evolution* **63**, 1-19 (2009).

58. Yang, Z. & Rannala, B. Molecular phylogenetics: principles and practice. *Nature Reviews Genetics* **13**, 303-314 (2012).

59. Dziedzioch, C., Stevens, A.D. & Gottsberger, G. The hummingbird plant community of a tropical montane rain forest in southern Ecuador. *Plant Biology* **5**, 331-337 (2003).

60. Cronk, Q. & Ojeda, I. Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* **59**, 715-727 (2008).

61. Colosimo, P.F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928-1933 (2005).

62. Kronforst, M.R. *et al.* Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *PNAS* **103**, 6575-6580 (2006).

63. Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* **455**, 620-626 (2008).

64. Marchinko, K.B. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* **63**, 127-138 (2009).

65. Stuurman, J. *et al.* Dissection of floral pollination syndromes in Petunia. *Genetics* **168**, 1585-1599 (2004).

66. Galliot, C., Hoballah, M., Kuhlemeier, C. & Stuurman, J. Genetics of flower size and nectar volume in *Petunia* pollination syndromes. *Planta* **225**, 203-212 (2006).

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

68. Bouck, A., Wessler, S.R. & Arnold, M.L. QTL analysis of floral traits in Louisiana Iris hybrids. *Evolution* **61**, 2308-2319 (2007).

69. Brothers, A.N. *et al.* Genetic architecture of floral traits in *Iris hexagona* and *Iris fulva*. *Journal of Heredity* **104**, 853-861 (2013).

70. Wessinger, C.A., Hileman, L.C. & Rausher, M.D. Identification of major quantitative trait loci underlying floral pollination syndrome divergence in *Penstemon*. *Philosophical Transactions of the Royal Society B* **369**, 20130349 (2014).

71. Bradshaw, H.D., Otto, K.G., Frewen, B.E., McKay, J.K. & Schemske, D.W. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (Mimulus). *Genetics* **149**, 367-382 (1998).

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

73. Charlesworth, D. & Charlesworth, B. Selection of recombination in clines. *Genetics* **91**, 581-589 (1979).

Figure Legends for main text

Figure 1 The hawkmoth pollinator, *Manduca sexta*, prefers ultra-violet (UV)-absorbent flowers to UV-reflective flowers. (a,b) The flowers of *P. axillaris* Rio Arapey (Arapey) and *P. axillaris* S7 (S7) are white in visible light (a) but only the Arapey accession absorbs UV-light (b). (c) *M. sexta* prefers Arapey over S7 on a first choice basis when presented with the two accessions in a flight cage. Binomial test, $p = 0.019$, $n = 20$. (d) The mean number of feeding events per moth by *M. sexta* was higher on Arapey than S7. Wilcoxon signed-rank test, $Z = -3.747$, $p < 0.0001$, $n = 20$. (e) Mean time spent feeding was equivalent for both accessions. Wilcoxon signed-rank test, $Z = -0.373$, $p = 0.73$, $n = 20$. In all graphs, bars show the mean±SE. Scale bar represents 2 cm.

Figure 2 A single major locus determines differences in UV-absorption between *P. inflata*, *P. axillaris* and *P. exserta*. (a) Flowers of *P. inflata*, *P. axillaris* and *P. exserta* and their F_1 in visible light and UV light with a schematic portraying the phylogenetic relationship of the three species. Pollination syndromes are

indicated and, at the nodes, the inferred ancestral pollination syndromes. (b) Flavonol content of the corolla limb of each species as measured by LC-MS. Bars show mean±SD, *n* = 3 except for the *P. axillaris* x *P. inflata* F_1 (*n* = 2). (c) QTL likelihood profile for UV-absorbance in the *P. axillaris* x *P. exserta* F_2 mapping population shows a major QTL on chromosome II that explains 79% of the variation in this trait (Supplementary Table 1). The seven chromosomes of *Petunia* are given on the x axis. The blue line represents the logarithm of odds (LOD) score; the horizontal lines correspond to the threshold for highly significant (*p* = 0.01, pink) and significant (*p* = 0.05, yellow) LODs determined by resampling. (d) A negative association is seen between UV- and visible-absorbance in the *P. axillaris* x *P. exserta* F_2 population. Spearman's rank-order correlation, $p < 0.0001$, *n* = 192. (e) Complementation tests of *P. inflata*, *P. hybrida* W138 and S7 with *P. exserta* give UV-reflective flowers, indicating they are all affected in the same locus required for high flavonols and thus UV-absorbent flowers. A UV-absorbent *P. axillaris* is included for comparison. Scale bars represent 2 cm.

Figure 3 *FLS* expression in floral buds of *P. axillaris*, *P. inflata* and *P. exserta*. (a) *FLS* expression is 8-fold higher in floral buds (stage 4; Supplementary Fig. 2b,c) of *P. axillaris* than *P. inflata.* (b) *FLS* expression is 56-fold higher in floral buds (stage 6; Supplementary Fig. 2b) of *P. axillaris* compared to *P. exserta*. (c) *FLS* expression in different developmental stages (Supplementary Fig. 2b) using quantitative RT-PCR is consistently higher in *P. axillaris* than *P. exserta*. Expression is normalised to stages 5/6 of one *P. axillaris* biological replicate. (d) *FLS* shows no significant difference in allele specific expression (ASE) in the *P. axillaris* x *P. inflata* F_1. (e) Expression of the *P. axillaris* allele is higher than the *P. exserta* allele in the introgression line, IL2-1, which is in the *P. exserta* genetic background and heterozygous for *P. axillaris* over the chromosome II QTL region ($IL2\text{-}1^{het}$)[37]. Unless otherwise indicated, expression was measured by RNA-seq. For (a,b,c), bars show mean±SD, *n* = 3. For (d), bars show mean±SD of reads mapped over 14 SNPs ($p = 0.48$; H_0 = no ASE, *n* = 3). For (e), bars show mean±SD of reads mapped over 1 SNP ($p < 0.01$; H_0 = no ASE, *n* = 4). For details of statistical tests in (d,e) see Online Methods.

Figure 4 Transposon insertions in the *MYB-FL* gene reduce flavonol levels and increase anthocyanin levels. (a,b) A flower from the *P. axillaris* x W138 F_1 individual A1-95 (sample A1-95_2 in Supplementary Table 2) shows a transposon-induced sector in UV light (a) and visible light (b). (c) Flavonol content, quantified by LC-MS, was lower in UV-reflective tissue than in UV-absorbent tissue (sample A1-95_3 in Supplementary Table 2). Bars shows mean±SD for *P. axillaris* (*n* = 3) and W138 (*n* = 2); UV-absorbent, UV-reflective tissue samples (*n* = 1). (d) Five UV-reflective somatic sectors of independent origin had transposon insertions in the *MYB-FL* gene. Labels on transposons insertions correspond to Supplementary Table 2. Filled boxes represent exons. (e) A complementation test of the germ-line insertion mutant A1-95 (genotype $MYB\text{-}FL^{Pax\text{-}dTph4}/MYB\text{-}FL^{W138}$) with *P. exserta* gives progeny that are all UV-reflective. This demonstrates that the $MYB\text{-}FL^{Pax\text{-}dTph4}$ allele is inactive. A single representative flower photographed in UV light is shown from each genotypic class and a *P. axillaris* flower is included for comparison. (f) Anthocyanidins (non-glycosylated anthocyanins) quantified in the same UV-absorbent and UV-reflective samples as (c). (g) *MYB-FL* expression during flower development (Supplementary Fig. 2b) is lower in *P. exserta* than *P. axillaris*. Expression normalised to stages 5/6 of one *P. axillaris* biological replicate. Bars show mean±SD, *n* = 3. (h) Partial *MYB-FL* sequence alignment showing the single base-pair deletion (orange arrow) in *P. exserta* that is predicted to lead to a frameshift and premature stop codon (blue arrow).

Figure 5 *MYB-FL* shows differential expression potentially due to sequence differences in the promoter and second intron. (a) *MYB-FL* shows 24-fold higher expression in floral buds (stage 4; Supplementary Fig. 2b,c) of *P. axillaris* than *P. inflata*. (b) *MYB-FL* is expressed predominantly from the *P. axillaris* allele in the *P. axillaris* x *P. inflata* F_1. (c) Schematic of *MYB-FL* gene sequences in *P. inflata* and *P. axillaris* indicating structural polymorphisms in the promoter (977 bp insertion) and second intron (254 bp *MITE* insertion in *P. axillaris* and *P. exserta*; approx. 715 bp insertion in *P. inflata* with homology to *dTph4* transposon). Filled boxes represent exons; numbers indicate order of exons; "T" indicates the location of the TATA-binding box. *MYB-FL* expression was measured by RNA-seq. In (a), bars show mean±SD, $n = 3$; in (b), bars show mean±SD reads mapped over 13 SNPs ($p < 0.0001$; H_0 = no allele-specific expression, $n = 3$) For details of statistical test in (b) see Online Methods.

Figure 6 UV-absorbance is associated with *MYB-FL* genotype in segregating populations. Progeny of putative hybrid accessions showing intermediate phenotypes were phenotyped for UV-absorbance and genotyped for *MYB-FL*. Plants 1 and 7 were sampled from the Ponto 143 population; the remaining plants were sampled from Pedra da Cruz (Supplementary table 5). Progeny that had at least $n = 4$ per genotypic class (progeny of plants 1, 2, and 3) were tested for statistically significant differences (Kruskal-Wallis one-way analysis of variance). Genotypes were obtained for three parental plants: plant 2 (H), plant 4 (A) and plant 6 (A). The box bounds the inter-quartile range (IQR) divided by the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Data beyond the end of the whiskers are outliers and plotted as points (as specified by Tukey). The width of the boxes are relative to *n* within a single plot. E, homozygous for single base-pair deletion, *P. exserta*-like allele; H, heterozygous; A, homozygous for *P. axillaris*-like allele.

Table 1 Contribution of the *P. axillaris* chromosome II QTL region to expression of flavonoid biosynthetic genes in *P. exserta*

							Normalised read counts	
Gene	Annotation in *P. axillaris* genome	Genbank reference	Present in introgression	Comparison	Fold change	Sig. level	*P. exserta*/IL2-1^{Pex}	IL2-1^{Pax}
FLS	Peaxi162Scf00927g00035	Z22543	Yes	IL2-1^{Pax} vs *P. exserta*	36.7	***	717	29,843
				IL2-1^{Pax} vs IL2-1^{Pex}	32.8	***	824	29,843
MYB-FL	Peaxi162Scf00886g00028	KT962949	Yes	IL2-1^{Pax} vs *P. exserta*	2.0	***	1,410	2,946
				IL2-1^{Pax} vs IL2-1^{Pex}	1.6	*	1,848	2,946
CHS-J	Peaxi162Scf00536g00092	X14597	No	IL2-1^{Pax} vs *P. exserta*	2.4	***	8,698	21,341
				IL2-1^{Pax} vs IL2-1^{Pex}	2.7	***	7,830	21,341
DFR-A	Peaxi162Scf00366g00630	X79723	No	IL2-1^{Pax} vs *P. exserta*	0.8	***	8,874	6,768
				IL2-1^{Pax} vs IL2-1^{Pex}	0.8	***	9,002	6,768
ANS	Peaxi162Scf00620g00533	X70786	No	IL2-1^{Pax} vs *P. exserta*	0.9	ns	15,911	14,028
				IL2-1^{Pax} vs IL2-1^{Pex}	0.9	ns	15,784	14,028
RT	Peaxi162Scf00487g00064	X71060	No	IL2-1^{Pax} vs *P. exserta*	0.7	*	28,639	20,370
				IL2-1^{Pax} vs IL2-1^{Pex}	0.7	**	29,485	20,370
MT	Peaxi162Scf00518g00430	KJ639934	No	IL2-1^{Pax} vs *P. exserta*	0.9	ns	15,748	14,893
				IL2-1^{Pax} vs IL2-1^{Pex}	0.9	*	17,270	14,893
GST	Peaxi162Scf00713g00038	Y07721	No	IL2-1^{Pax} vs *P. exserta*	0.8	ns	8,088	6,356
				IL2-1^{Pax} vs IL2-1^{Pex}	0.7	**	8,997	6,356

A comparison of gene expression in mature buds between the chromosome II introgression line IL2-1^{Pax} and *P. exserta*, in order to determine the effect of the chromosome II QTL locus on flavonoid biosynthetic genes. IL2-1 is in the *P. exserta* genetic background, with a region on chromosome II homozygous for *P. axillaris* (IL2-1^{Pax}) or *P. exserta* (IL2-1^{Pex}). The comparison between IL2-1^{Pax} and IL2-1^{Pex} is included to assess the contribution of genomic regions still potentially segregating in the background of the introgression line. *FLS*, *flavonol synthase*; *CHS-J*, *chalcone synthase J*; *DFR-A*, *dihydroflavonol reductase A*; *ANS*, *anthocyanidin synthase*; *RT*, *rhamnosyltransferase*; *MT*, *methylation at three*; *GST*, *glutathione S-transferase*; Sig. level, *p* statistic calculated using DESeq2 and adjusted for multiple testing with the Benjamini-Hochberg procedure to control for false discovery rate; ns, not significant; *$p < 0.05$, **$p < 0.01$, ***$p < 0.0001$.

Online Methods

Plant Material

Petunia axillaris ssp *axillaris* N (*P. axillaris*) is from the Rostock Botanical Garden (Germany); *P. a. ssp. parodii* S7 (S7) and *P. inflata* S6 (also referred to as *P. integrifolia* ssp. *inflata* S6*; P. inflata*) were provided by R. Koes (University of Amsterdam, the Netherlands); and *P. exserta* from R. J. Griesbach (Beltsville, USA). The accessions are maintained by cuttings and self-fertilization, except for *P. inflata* which is maintained by sibling crosses. *P. a.* ssp *parodii* Rio Arapey was collected as selfed seeds in Uruguay in January 2006 (S 30° 58' 09.5'' W 57° 41' 38.1''). *P. hybrida* W138 is a standard laboratory line maintained by self-fertilization. The introgression line, IL2-1, results from a *P. exserta x P. axillaris* F_1 backcrossed to *P. exserta* five times and selfed three times, and is further described in ref. 37. Wild accessions of *P. axillaris* and *P. exserta* were sampled over multiple field seasons (2002, 2003, 2006, 2007/2008 for *P. axillaris*; 2002, 2007/2008 for *P. exserta*) and are described further in Supplementary Table 5 and in ref. 49. Progeny of individuals showing intermediate floral traits come from natural pollinations collected from individuals at the sympatric populations, Pedra da Cruz and Ponto 143 (Supplementary Table 5), on field trips in November 2013. For hybrid individuals 2-6, 8 and 9, the total amount of seeds collected were sowed and for individuals 1 and 7, 46 and 24 seeds were sowed, respectively; the progeny represent the total number of seeds that germinated and reached maturity. Plants were grown as described[13]. Plants with tissue sampled for liquid chromatography-mass spectrometry (LC-MS) and proton transfer reaction mass spectrometry (PTR-MS) analyses, quantitative RT-PCR, and RNA-seq experiments were grown in a growth chamber under a light:dark regime of 15:9 h at 22 °C (day) and 17 °C (night), and tissue was sampled at the same time each day during a single experiment.

Behavioral experiments

M. sexta (strain Yamamoto) pupae were obtained from the North Carolina State University insectary and held under the following climatic conditions: light was provided by six fluorescent tubes (Philips TDL, 36 W, >1 kHz) with a 16:8 L:D cycle in an environmental cabinet programmed at 26 °C and 65% relative humidity. Pupae were sexed prior to emergence and placed in rearing cages (BugDorm-4180F; BugDorm) containing a wet tissue. Naïve females were used for the experiments and were 4-6 days old, unmated and unfed. Experiments were conducted in a screen cage (368 x 248 x 144 cm) outside in the Botanical Garden of Bern (Switzerland) during October 2007, and August/September, 2008. The tests were performed between 30 min to 2 h after sunset, essentially as described previously[41]. Naïve moths were taken one by one and released in the experimental cage for 5 min. Each moth was presented with two plants of each species, each with equal numbers of flowers. The position of the plants was alternated day to day. The entrance site of the moths varied between each assay, to minimize the impact of the first flower seen on entry. The moths could fly freely and during this time their behavior bouts (feeding and duration of feeding events) were recorded. The number of moths tested was based on previous behavioral experiments[41]. Moths that did not feed were discarded.

UV images and UV scoring

Images were recorded using a Nikon 60 mm 2.8D micro lens with a Nikon D7000 SLR camera that was converted to record UV light by replacing the manufacturer's filter with a UV-specific filter that blocks visible and infra-red light (Advanced Camera Services Ltd). A light source was provided by a Metz MZ76 flash gun that was modified to produce UV-A light (320-390 nm; Advanced Camera Services Ltd). Images were converted to greyscale in Photoshop CS4 (Adobe Systems) and, where necessary, exposure was adjusted over the complete image. Flowers were scored either as UV-absorbent or UV-reflective based on comparison with a representative *P. axillaris* flower.

Spectrophotometric quantification of flavonols and anthocyanins

For large tissue samples (e.g. measurement of flowers from F_2 populations), an 8 mm diameter disc of floral tissue was sampled from the corolla limb and put into 1 mL of extraction buffer (2:1:7, methanol:acetic acid:water). For small tissue samples (e.g. transposon tagging F_1 plants), a 4 mm diameter disc of tissue was placed into 30 µL of extraction buffer. After 48 h in the dark, absorption spectra were measured on a spectrophotometer (large samples: Ultraspec 3100 pro, GE Healthcare Life Sciences; small samples: Nanodrop ND-1000, Thermo Fisher). Flavonol values represent summed absorbance values over 315-378 nm. Anthocyanin values represent summed absorbance values over 445-595 nm, except for the Nanodrop where absorbances were summed according to the closest wavelength to these ranges. Comparisons were made between plants grown in the same growth conditions (i.e. equivalent exposure to light).

LC-MS quantification of flavonols and anthocyanidins

Tissue was collected from 1-2 day old flowers, homogenized in liquid nitrogen and stored at −80 °C until extraction. For experiment one (*P. axillaris*, *P. axillaris* x *P. exserta* F_1, *P. exserta* and A1-95 comparisons), corolla limb tissue from three flowers from the same plant constituted a single biological replicate; for experiment two (*P. axillaris* x *P. inflata* F_1 and *P. axillaris* comparison), corolla limb tissue of one flower of one plant represented one biological replicate. Flavonoids were extracted using 20 µL of extraction buffer (5% w/v hydrochloric acid in methanol, plus 50 nM of an internal standard) per 1 mg ground plant tissue. Samples were sonicated for 15 min following extraction and then hydrolyzed at 100 °C for 10 min in order to obtain the aglycones. Samples were centrifuged at 13,000 rpm and the supernatant was diluted (1:1,000) with methanol containing 1% HCl prior to LC-MS analysis. LC-MS analysis was performed using an Agilent 1260 HPLC-System connected to a QTRAP 5500 mass spectrometer with a Turbo V Source (ABSciex). Peak areas of each compound for each sample were normalized to the internal standard and quantification was done by external calibration to standards (kaempferol, quercetin, myricetin, cyanidin, malvidin; Sigma-Aldrich; petunidin, peonidin, Extrasynthese). Pelargonidin and delphinidin were too low to measure accurately so were excluded from analysis. Data for Figure 2b was combined from experiments one and two, using quantification relative to *P. axillaris.* In some samples (Fig. 2b: *P. axillaris*, *P. axillaris* x *P. exserta* F_1; Fig. 4c: *P. axillaris* and A1-95 UV-absorbent tissue), quercetin values were above the range of the standard curve (4.1x to 9.5x) so these concentrations were derived by extrapolation.

Measurement of additional floral traits

Floral surface area was recorded using front-view photographs and measured using ImageJ as described[42]. Nectar was weighed after extracting from corolla tubes by centrifugation. Methylbenzoate was measured using PTR-MS as described previously[13].

DNA/RNA manipulations

Genomic DNA extraction was performed with a modified CTAB method[74] or with the plant GenElute DNA isolation kit (Sigma), except for genomic DNA from wild accessions which was performed as described previously[49]. For RNA extractions, all tissue was homogenized in liquid nitrogen. RNA extracted for the RNA-seq experiment comparing *P. axillaris*, *P. exserta* and the IL2-1 breeding line was done using Trizol (Life Technologies) according to the manufacturer's recommendations. All other RNA extraction was carried out using the RNeasy Plant Mini Kit (Qiagen). DNA and RNA were quantified using a Nanodrop ND-1000 (Thermo Fisher). Oligonucleotides (Supplementary Table 6) were designed using Primer3.

Genotyping and marker association analyses

Three flowers from each of the 203 *P. axillaris* x *P. exserta* F_2 individuals were quantified for UV- and visible-absorbance using spectrophotometry. Genotyping and QTL analysis of this population is described elsewhere[13]: the QTL likelihood profile in Figure 2c was created using Qgene (simple interval mapping) and the statistics in Supplementary Table 1 were calculated using QTX[75]. Further genotyping was performed with the MYB-FL-5'CAPS marker: details of the *MYB-FL* CAPS markers and PCR conditions can be found on our website. A Spearman's rank correlation coefficient was used to test the association between UV- and visible-absorption in the *P. axillaris* x *P. exserta* F_2 population because the data most resembled a monotonic relationship. Wild accessions were genotyped using the MYB-FL-3'CAPS marker and some were screened using positive control PCRs (see Supplementary Note). Progeny of putative hybrid wild accessions were genotyped in the same way and phenotyped using spectrophotometry (five flowers per plant). Genotypic classes had equal variance (Levene's test) but were not normally distributed (Shapiro-Wilk test) so tests for association of means with *MYB-FL* genotype were carried out using the Kruskal-Wallace one-way analysis of variance in R.

Transposon tagging

F_1 populations were made by crossing *P. hybrida* W138 red flower (*AN1*-revertant) pollen to both *P. axillaris* and IL2-1^{Pax}. We screened for sectors of interest initially using a UV camera: three flowers per plant from 80 plants of the *P. axillaris* x W138 F_1 population and three flowers per plant from 64 plants of the IL2-1^{Pax} x W138 F_1 population. A total of 28 UV-reflective sectors were obtained ranging in size from 2 mm^2 to 75 mm^2. These sectors fit four major classes based on appearance in visible light: white (5), dark purple (12), bright pink (9), and the same color as the surrounding tissue (2). The bright pink class reflected the most UV light in comparison to the surrounding F_1 tissue (Fig. 4a,b). We therefore focused on identifying large sectors of this bright pink class, allowing further screening to be done by eye. Over 14 months, we screened the constantly flowering F_1 populations twice weekly and sampled bright pink sectors that were at least 100 mm^2 ($1/10^{th}$ petal for the *P. axillaris* x W138 F_1 and $1/8^{th}$ petal for the IL2-1^{Pax} x W138 F_1). In total, 45 of such large sectors and neighboring tissue were sampled from 36 individual plants. Small tissue samples from the sector and surrounding tissue were taken for spectrophotometric

analysis, and tissues were separated and frozen in liquid nitrogen. DNA was extracted from both tissues and PCR screened. A *P. axillaris* transcriptome assembly from stage 5/6 floral bud tissue was used to blast for homologues to the *Arabidopsis* genes *AtMYB111* (Genbank NM_124310), *AtMYB12* (Genbank NM_130314) and *AtMYB11* (Genbank NM_116126). A contig corresponding to the first two exons of the *AtMYB111* gene was found and designated *MYB-FL*, and primers were designed to the first two exons to use in the PCR screen (B371 and B369 primers; Supplementary Table 6). Samples that showed a difference in *MYB-FL* PCR product size between the sector and surrounding tissue were scored as potentially containing a transposon insertion (Supplementary Table 2). Some of these sector samples also showed the *MYB-FL*-sized PCR product presumably due to amplification of non-mutated *MYB-FL* in the L2 and L3 tissue layers. Products were cloned using the pGEM®-T Easy vector system (Promega) and subsequently sequenced.

A transposon display protocol[76], modified to allow sequencing by Illumina technology on a High Seq2000 platform (Baseclear, The Netherlands) was also carried out on sixteen samples. To guarantee a broad coverage of insertion loci, we independently amplified from each sample three possible flanking sequence collections, using either a MseI/MfeI digestion (recovery of *dTph1* right border flanking sequences) or a BfaI/MfeI digestion (recovery of *dTph1* left and right border flanking sequences). The origin of the different samples in the sequencing output was determined based on the incorporation of barcodes in the amplification primers. The resulting sequences were blast-searched for novel insertions into *MYB-FL* that were present in material derived from sectors, and absent from surrounding tissue. For further details see Supplementary Note.

To obtain germline mutations, tissue culture was carried out on the sepals of a fully UV-reflective flower from the *P. axillaris* x W138 F_1 plant, A1-95 (sample A1-95_3 in Supplementary Table 2). Resulting clonal plants were selfed and 24 progeny per plant were phenotyped for UV-absorbance (fully UV-reflective progeny indicated the *MYB-FL*$^{Pax\text{-}dTph4}$ allele was present in the germ-line), and PCR screened for the *MYB-FL* allele (B371 and B369 primers; see Supplementary Table 6). This germline mutant was crossed with *P. exserta* (*MYB-FL*^{Pex}/*MYB-FL*^{Pex}) to give progeny of genotype *MYB-FL*^{Pex}/*MYB-FL*$^{Pax\text{-}dTph4}$ *or MYB-FL*^{Pex}/*MYB-FL*^{W138}. These plants were phenotyped and genotyped as described above.

Sequence analysis of *MYB-FL*

The full-length *MYB-FL* sequences including promoter, were obtained by blasting the draft genomes of *P. axillaris* (v.1.6.2) and *P. inflata* (v.1.0.2) and were confirmed by Sanger sequencing (see primers in Supplementary Table 6). The sequence of *MYB-FL* in *P. exserta* was obtained using PCR amplification with primers designed to the *P. axillaris* sequence and Sanger sequencing (see primers in Supplementary Table 6). Sequence analyses were performed in Geneious (Biomatters).

Quantitative RT-PCR

Tissue was collected from different floral developmental stages (Supplementary Fig. 2b) and samples consisted of: stage 1, three buds, sepals removed; stage 2, two buds, sepals and reproductive organs removed; stage 3 and 4, 1 bud, sepals and reproductive organs removed; stages 5/6 and 7/8, 1 bud each, sepals and reproductive organs removed. Bud tissue was combined during homogenization of the tissue.

Samples were collected from three plants representing three biological replicates. After extraction, RNA samples were treated with DNAse I (Sigma-Aldrich) and the quality of the RNA subsequently measured on a 2100 Bioanalyzer (Agilent). Samples with RNA Integrity Numbers (RIN) less than 7.5 were discarded. First strand synthesis was performed using Transcriptor Universal cDNA Master (Roche) containing random hexamer primers, according to the manufacturer's recommendations. Quantitative RT-PCR experiments were performed using a LightCycler® 96 Real-Time PCR System (Roche) with KAPA SYBR® FAST qPCR Kit optimized for LightCycler® 480 (KAPA Biosystems), according to the manufacturer's recommendations. Reactions were run in triplicates for each gene. Primers are given in Supplementary Table 6. Cycle of quantification (Cq) thresholds and normalization calculations were determined by LightCycler® 480 Software (v.1.1.0.1320; Roche). Each biological replicate of each species was analyzed on a separate PCR plate and a single *P. axillaris* stage 5/6 sample from one biological replicate was included in each plate as a normaliser. Integration of data from biological replicates was conducted manually. Standard curves were performed to determine the PCR efficiencies. No reverse-transcription controls were performed for each sample. Reference genes[77] were tested for stability of expression on two developmental stages from each species using NormFinder[78]. *SAND* and *RAN1* were shown to be the most stable reference genes and were included for each sample.

RNA sequencing

For experiment one (comparisons of *P. axillaris*, *P. exserta* and IL2-1), three stage 5/6 buds (sepals and reproductive organs removed; Supplementary Fig. 2b) from a single plant were combined for one biological replicate. Three biological replicates of *P. axillaris* and *P. exserta*, and four biological replicates of IL2-1^{Pax}, IL2-1^{het} and IL2-1^{Pex} were sequenced. For experiment two (comparisons of *P. axillaris*, *P. inflata* and *P. axillaris* x *P. inflata* F_1), three biological replicates were used per sample, and each biological replicate represented six stage 4 buds (sepals and reproductive organs removed; Supplementary Fig. 2b,c) from a single plant. The sepals and reproductive organs were removed, the corolla was dissected into limb and tube, and the limb tissue used for RNA extraction.

RNA was sent to the Lausanne Genomic Technologies Facility (Lausanne, Switzerland) for sequencing. Quality of RNA was checked using a Fragment Analyzer (Advanced Analytical). For experiment one, RNA Quality Numbers (RQN) ranged from 6.4 to 8.1; for experiment two, RQN ranged from 8.3 to 9.2. Cluster generation was performed with the sequencing libraries using the Illumina TruSeq PE Cluster Kit (v.3). For experiment one, samples were paired-end sequenced for 50 bp; for experiment two, samples were single-end sequenced for 100 bp. Sequencing data were processed using the Illumina Pipeline Software v.1.82. For all samples, raw reads were checked for contamination by aligning them against rRNA sequences from *P. axillaris*, the *Escherichia coli* genome and the human transcriptome using bowtie2 (v.2.2.1)[79]. Reads aligning to mentioned sequences were discarded. FastqMcf (v.1.1.2-686) was used to remove Illumina adapter sequences and trim low quality regions. After trimming, reads shorter than 40 bp and 60 bp were discarded for reads with initial length of 50 bp and 100 bp, respectively. These pre-processed reads were mapped against the draft reference genome of *P. axillaris* (v.1.6.2) using the SNP-tolerant and splice-aware aligner GSNAP (v.2015-06-12)[80]. To allow for variation between reference and reads, the "–m" option was set to 0.04. Variant calling was carried out in GATK (v.3.4.0)[81] according to GATK best practices for RNA-seq data. After duplicate marking and splitting reads with *N* in their CIGAR

string, local realignment around indels was undertaken, and base quality scores were recalibrated, using a set of high quality SNPs determined by an initial run of the GATK HaplotypeCaller. Only positions which showed homozygous and biallelic SNPs between parental species were considered to detect allele-specific expression, and SNPs included for consideration were in the exon, intron, 5'UTR (within 100 bp of start codon) and 3'UTR (within 100 bp of stop codon) regions. Total RNA-seq reads from all *P. axillaris* samples from experiment one were assembled with Trinity[82]. Trinity was run with quality trimming (-trimmomatic) and strand-specificity (-SS_lib_type RF) enabled.

Differential expression analyses

Reads mapping to the genes of interest were counted using HTseq (v.0.6.1)[83] in the union mode. Differential expression analysis was performed in R using DESeq2 (v.1.8.1)[84] with the default parameters, including the Cook's distance treatment to remove outliers. Normalized read counts are presented in figures. Allelic coverage for variant positions was detected with ASEReadCounter implemented in GATK[81,85]. Analyses of allelic imbalance were conducted in R with the package MBASED (v.1.2.0)[86] using rounded read counts averaged over the biological replicates. Parameters of read count overdispersion were estimated with a custom R script provided by the author of the MBASED package (available on request). The parameter *rho* was estimated as 0.00277 and 0.00145 using a set of 111 and 1766 genes for experiment 1 (IL2-1^{Pax} cf. IL2-1^{Pex}) and experiment 2 (*P. axillaris* cf. *P inflata*), respectively. Genes used to estimate *rho* had only 1 SNP, had minor expression changes between alleles (0.7-1.4x difference), and varied their major allele over the replicates.

Methods only references

74. Murray, M.G. & Thompson, W.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**, 4321-4325 (1980).

75. Manly, K.F., Cudmore, R.H. & Meer, J.M. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome* **12**, 930-932 (2001).

76. Vandenbussche, M., Zethof, J. & Gerats, T. in *Plant Transposable Elements. Methods and Protocols*, (ed. Peterson, T.) 239-250 (Humana Press, 2013).

77. Mallona, I., Lischewski, S., Weiss, J., Hause, B. & Egea-Cortines, M. Validation of reference genes for quantitative real-time PCR during leaf and flower development in *Petunia hybrida*. *BMC Plant Biology* **10**, 4 (2010).

78. Andersen, C.L., Jensen, J.L. & Ørntoft, T.F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research* **64**, 5245-5250 (2004).

79. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).

80. Wu, T.D. & Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873-881 (2010).

81. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491-498 (2011).

82. Grabherr, M.G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 644-652 (2011).

83. Anders, S., Pyl, P.T. & Huber, W. HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2014).

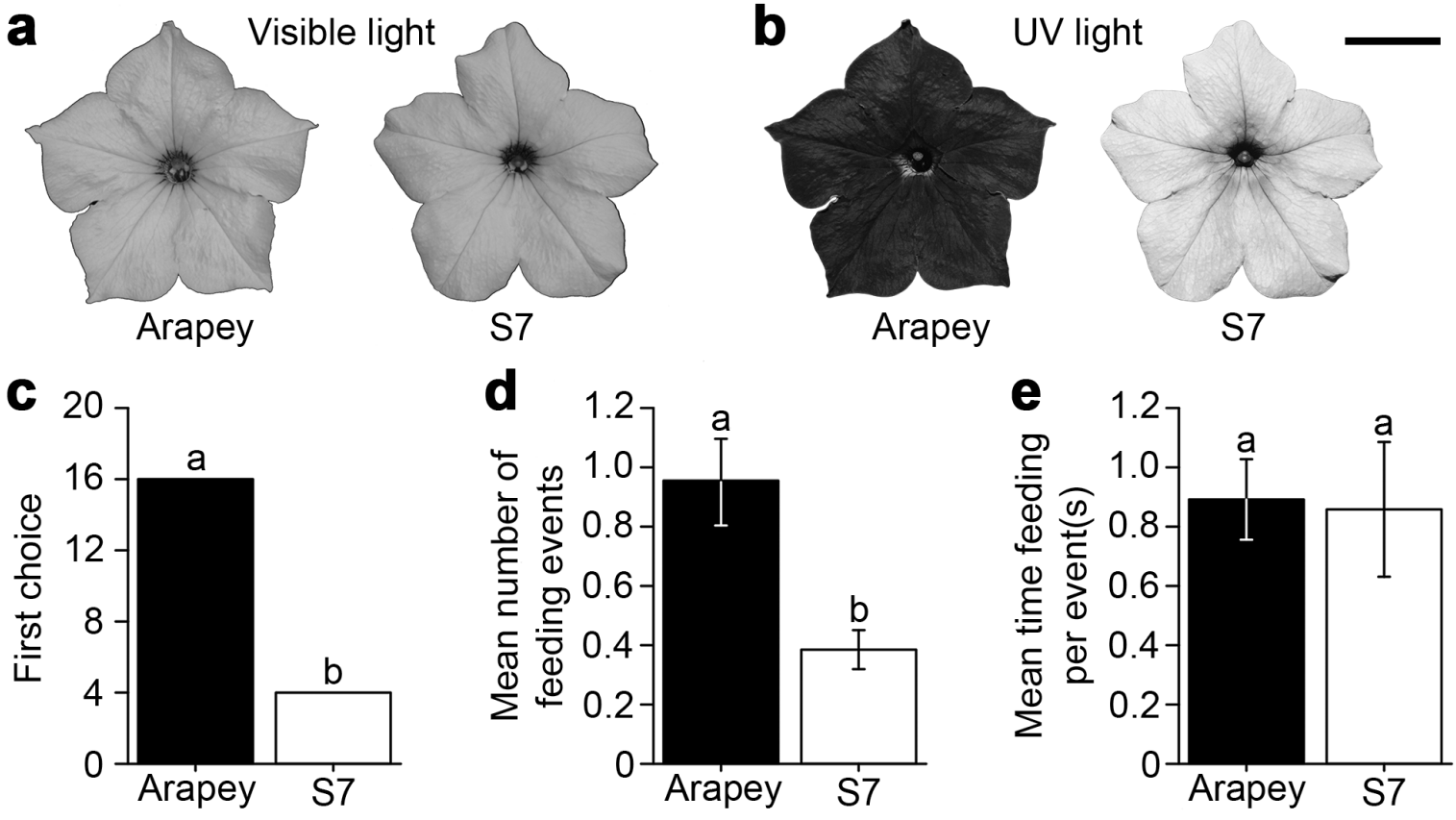
84. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

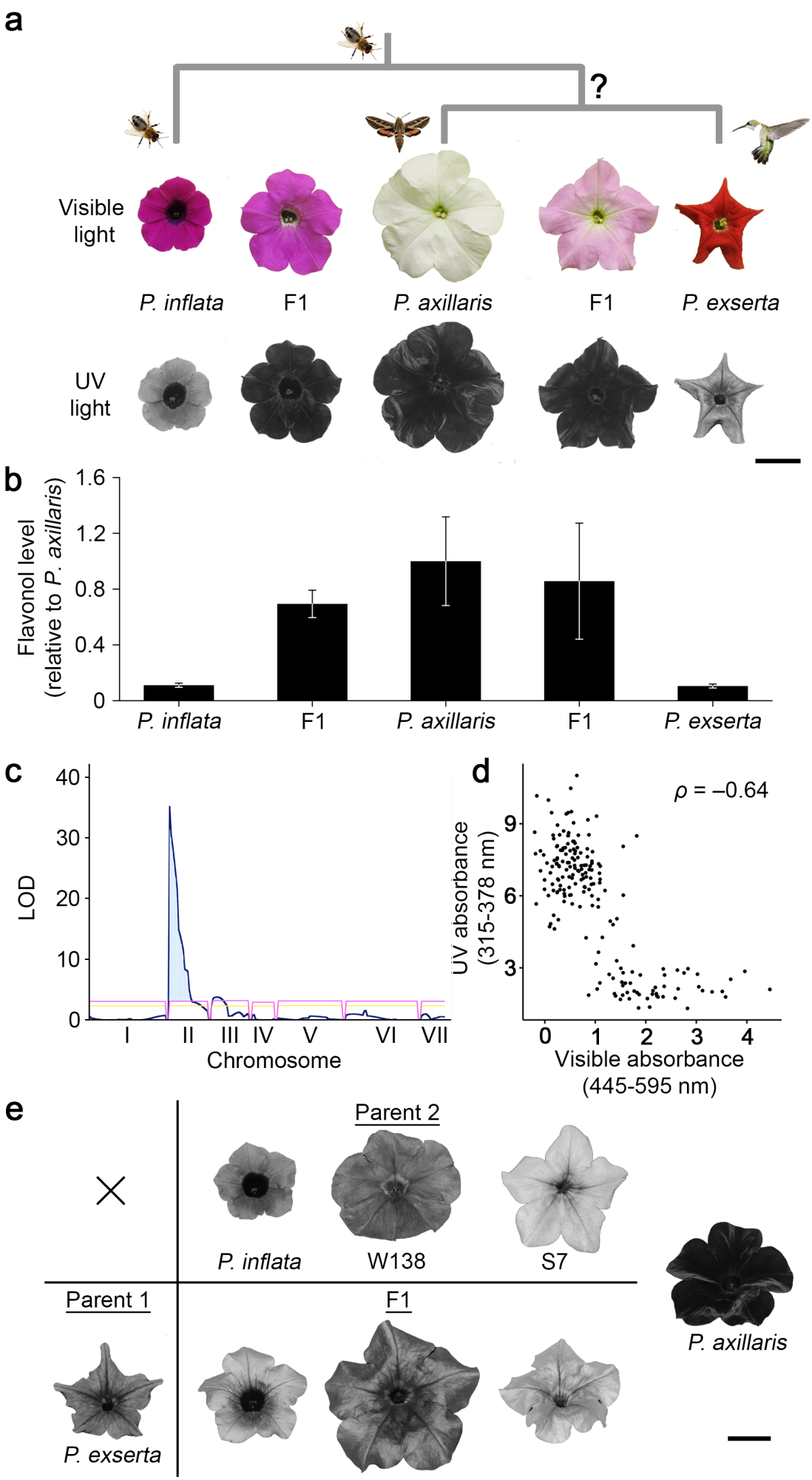
85. Castel, S.E., Levy Moonshine, A., Mohammadi, P., Banks, E. & Lappalainen, T. Tools and best practices for data processing in allelic expression analysis. *Genome Biology*, **16**, 195 (2015).

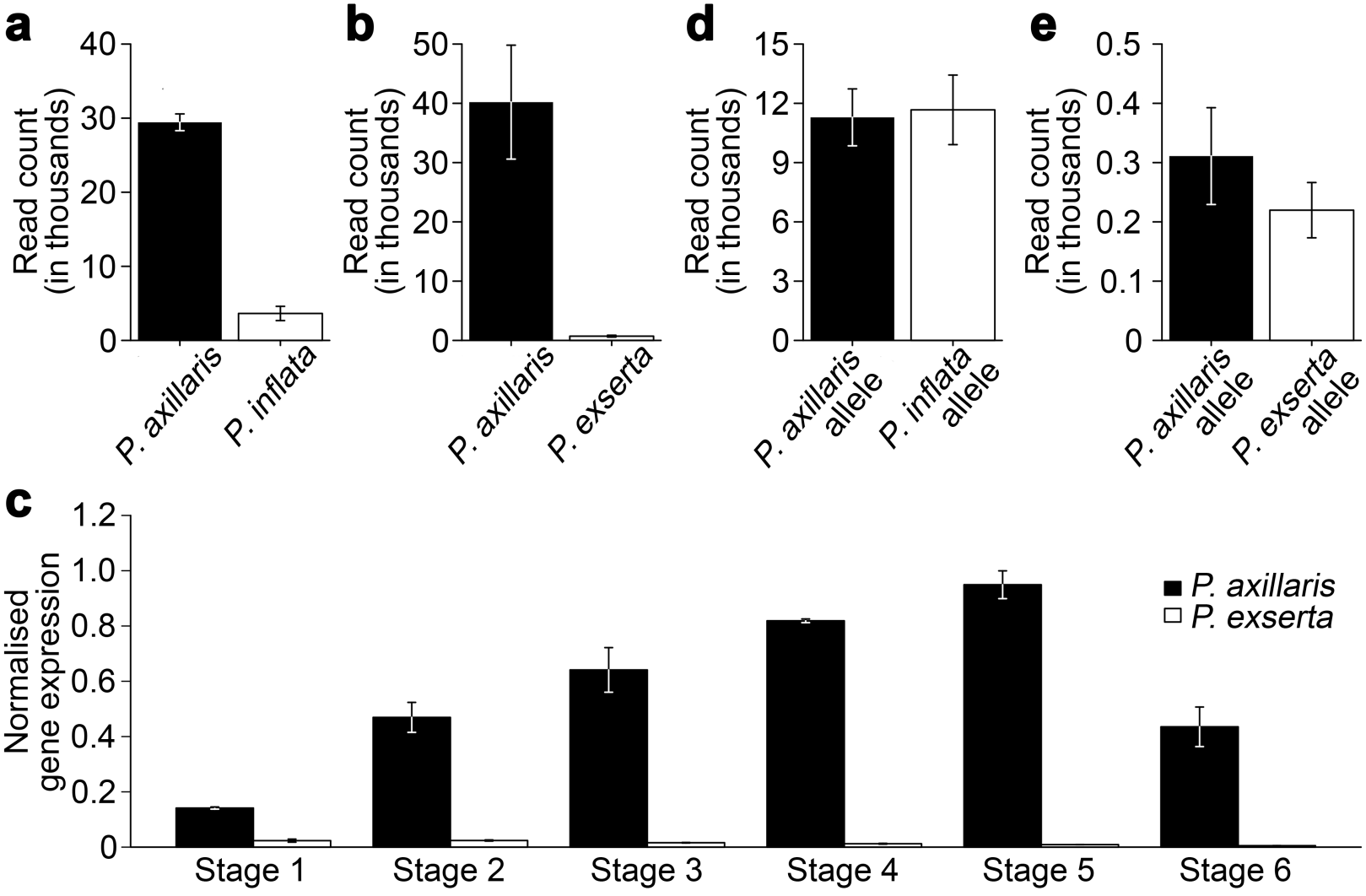
86. Mayba, O. *et al.* MBASED: allele-specific expression detection in cancer tissues and cell lines. *Genome Biology* **15**, 405 (2014).

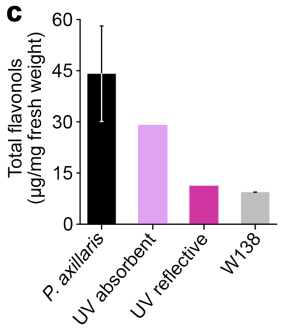
Competing Financial Interests

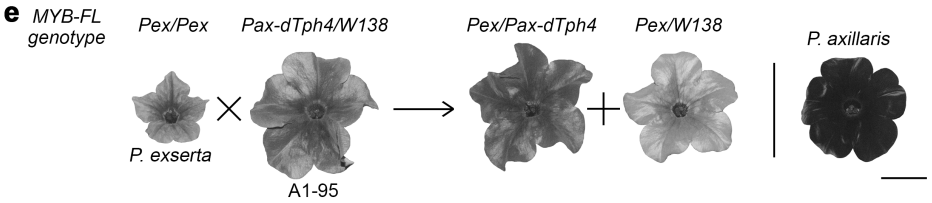
The authors declare no competing financial interests.

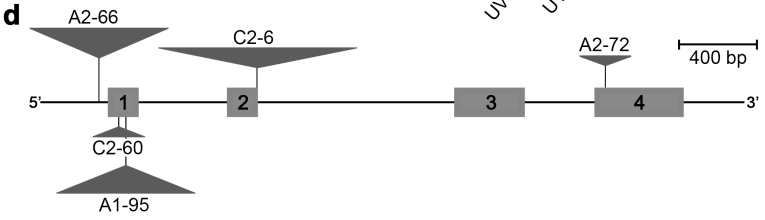

a
Visible light
Arapey
S7
b
UV light
Arapey
S7
c
First choice
20
16
12
8
4
0
a
b
Arapey
S7
d
Mean number of feeding events
1.2
1.0
0.8
0.6
0.4
0.2
0
a
b
Arapey
S7
e
Mean time feeding per event(s)
1.2
1.0
0.8
0.6
0.4
0.2
0
a
a
Arapey
S7

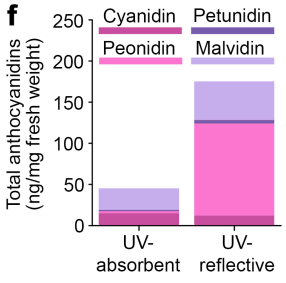

a
?
Visible light
P. inflata
F1
P. axillaris
F1
P. exserta
UV light
b
Flavonol level (relative to P. axillaris)
1.6
1.2
0.8
0.4
0
P. inflata
F1
P. axillaris
F1
P. exserta
c
LOD
40
30
20
10
0
I
II
III
IV
V
VI
VII
Chromosome
d
ρ = –0.64
UV absorbance (315-378 nm)
9
6
3
0
1
2
3
4
Visible absorbance (445-595 nm)
e
Parent 2
P. inflata
W138
S7
Parent 1
F1
P. exserta
P. axillaris

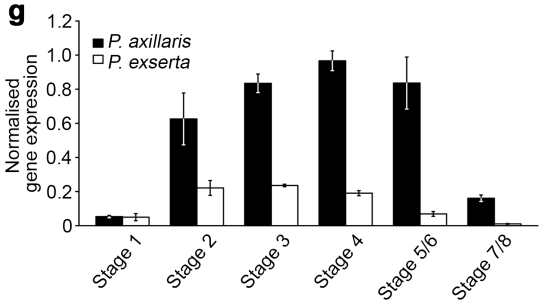

a
Read count (in thousands)
40
30
20
10
0
P. axillaris
P. inflata
b
Read count (in thousands)
50
40
30
20
10
0
P. axillaris
P. exserta
d
Read count (in thousands)
15
12
9
6
3
0
P. axillaris allele
P. inflata allele
e
Read count (in thousands)
0.5
0.4
0.3
0.2
0.1
0
P. axillaris allele
P. exserta allele
c
Normalised gene expression
1.2
1.0
0.8
0.6
0.4
0.2
0
P. axillaris
P. exserta
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Stage 6

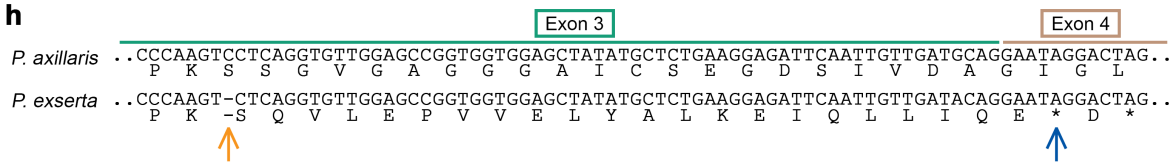

c
Total flavonols (µg/mg fresh weight)
60
45
30
15
0
P. axillaris
UV absorbent
UV reflective
W138

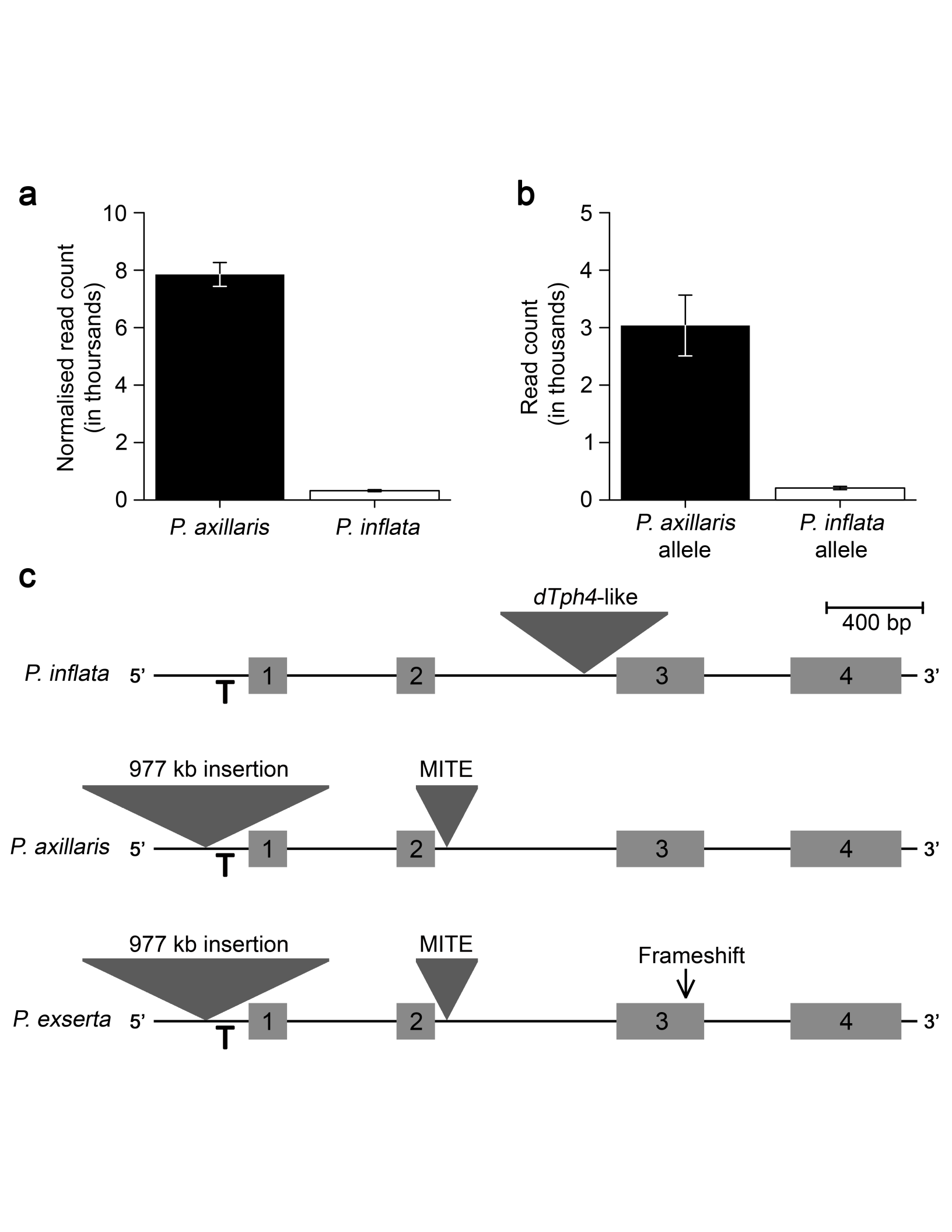

e
MYB-FL genotype
Pex/Pex
Pax-dTph4/W138
Pex/Pax-dTph4
Pex/W138
P. axillaris
P. exserta
A1-95

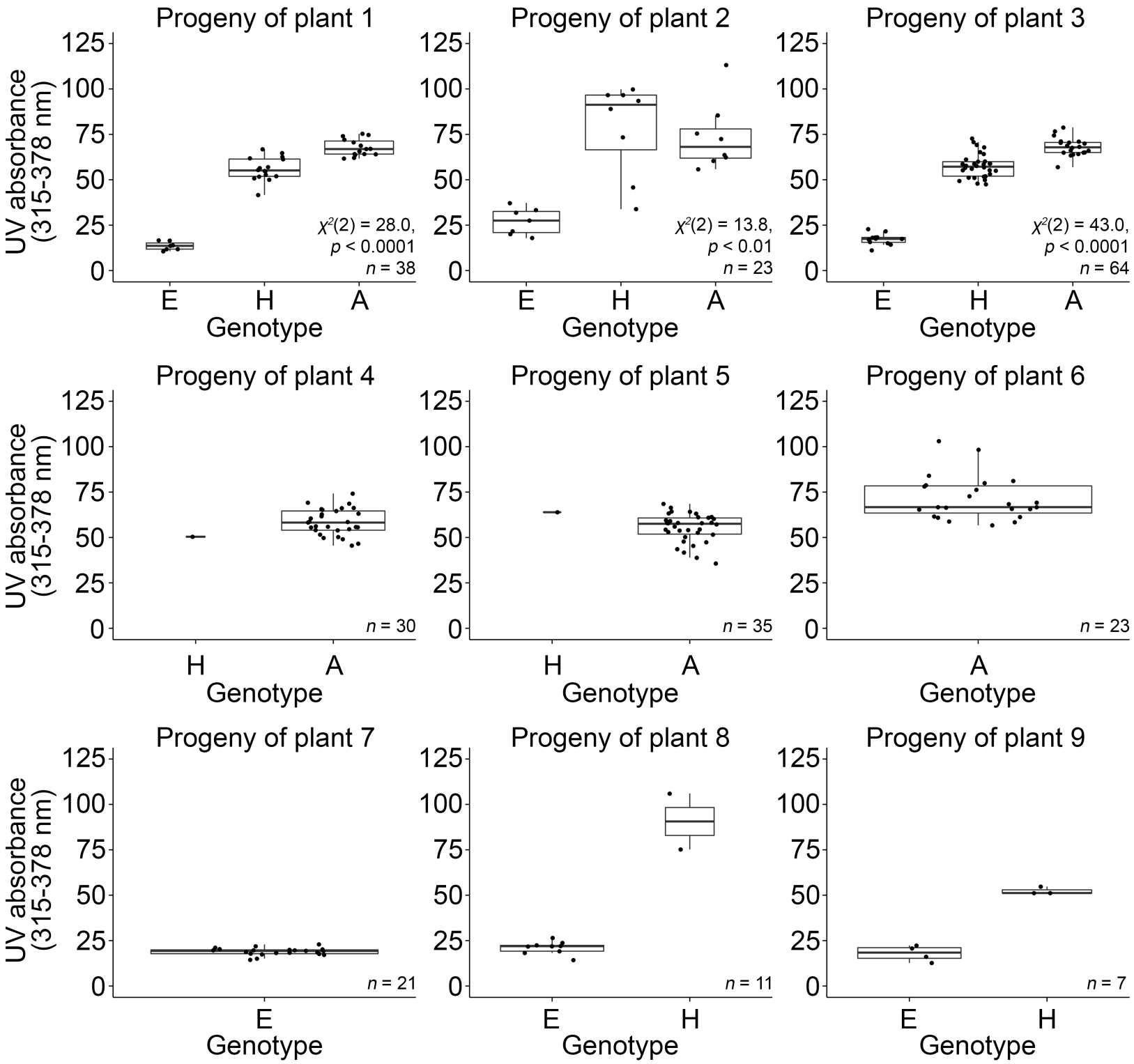

d
A2-66
C2-6
A2-72
400 bp
5'
1
2
3
4
3'
C2-60
A1-95


f
Total anthocyanidins (ng/mg fresh weight)
250
200
150
100
50
0
Cyanidin
Petunidin
Peonidin
Malvidin
UV-absorbent
UV-reflective


g
Normalised gene expression
1.2
1.0
0.8
0.6
0.4
0.2
0
P. axillaris
P. exserta
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5/6
Stage 7/8


h
Exon 3
Exon 4
P. axillaris ..CCCAAGTCCTCAGGTGTTGGAGCCGGTGGTGGAGCTATATGCTCTGAAGGAGATTCAATTGTTGATGCAGGAATAGGACTAG..
P K S S G V G A G G G A I C S E G D S I V D A G I G L
P. exserta ..CCCAAGT-CTCAGGTGTTGGAGCCGGTGGTGGAGCTATATGCTCTGAAGGAGATTCAATTGTTGATACAGGAATAGGACTAG..
P K -S Q V L E P V V E L Y A L K E I Q L L I Q E * D *


a
Normalised read count (in thoursands)
10
8
6
4
2
0
P. axillaris
P. inflata
b
Read count (in thousands)
5
4
3
2
1
0
P. axillaris allele
P. inflata allele
c
dTph4-like
400 bp
P. inflata
5'
T
1
2
3
4
3'
977 kb insertion
MITE
P. axillaris
5'
T
1
2
3
4
3'
977 kb insertion
MITE
Frameshift
P. exserta
5'
T
1
2
3
4
3'


Progeny of plant 1
Progeny of plant 2
Progeny of plant 3
Progeny of plant 4
Progeny of plant 5
Progeny of plant 6
Progeny of plant 7
Progeny of plant 8
Progeny of plant 9
UV absorbance (315-378 nm)
125
100
75
50
25
0
E
H
A
Genotype
χ²(2) = 28.0, p < 0.0001 n = 38
χ²(2) = 13.8, p < 0.01 n = 23
χ²(2) = 43.0, p < 0.0001 n = 64
n = 30
n = 35
n = 23
n = 21
n = 11
n = 7