THE MOLECULAR EVOLUTION OF FLORAL COLOUR SHIFTS IN *LATHYRUS* L. (FABACEAE)

by

Xinxin Xue

M.Sc., The University of Edinburgh, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in
THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
December 2015

Abstract

In this thesis, I used *Lathyrus* to study floral colour evolution. My work builds on that of pioneer geneticists who used the sweet pea (*Lathyrus odoratus*) for research in early 20th century England. They used the rich horticultural resource of inbred cultivars for genetic study. More recently, studies on flavonoid biochemistry and the phylogenetics of *Lathyrus* lend tools to place the colour shifts in an evolutionary and molecular context. In chapter 2, I show that the *A1* locus is a missense mutation (332 G/A) in the substrate recognition site (SRS1) of flavonoid 3',5'-hydroxylase (*F3'5'H*) and is associated with the pink mutant cultivar 'Painted Lady' ('PL'). This single base pair substitution in the mutant *F3'5'H* is speculated to toggle the enzyme from primary F3'5'H activity to a relatively efficient F3'H, as shown in a heterologous transformation in *Arabidopsis PAP1D* (a mutant line that produces anthocyanin constitutively). In chapter 3, I constructed a multi-species coalescent tree using Bayesian inference and reconstructed the ancestral states for floral colour, life history trait (perennial or annual) and floral pattern. The ancestral states for *Lathyrus* are anthocyanin rich (AR), annual and concolourous. However, no correlation was found between the life history trait (which is linked to breeding systems) and the loss of anthocyanin colour in the petal, when corrected for phylogenetic independence. This suggests that in *Lathyrus*, autogamous species are as colourful as allogamous ones even though the latter are expected to need greater floral display. In chapter 4, I found that the lack of expression in *dihydroflavonol reductase* (*DFR*) was associated with a white mutant cultivar 'Mrs Collier' ('MC') of *L. odoratus* via a *trans*-regulatory machinery. Two transcription factors, the sweet pea orthologues of *AN2* (MYB) and *AN1* (bHLH) were also not expressed although neither was associated with the white phenotype in an F2-cosegregation analysis. This *DFR* silencing was also observed in another white mutant of the domesticated grass pea (*L. sativus*). In contrast, when unpigmented wild species originate under natural selection, *DFR* expression, if at all affected, is lowered rather than fully silenced, likely due to pleiotropic effects.

Preface

This thesis represents original unpublished work by Xinxin Xue, with inputs from others as stated in the text and here. All chapters will be prepared for publication.

Chapter 2

I designed and carried out all the experiments and data analysis with inputs from others. Quentin Cronk provided the idea, performed the crosses and supplied F2 seed. Lina Madilao at the Bohlmann lab carried out the LC analysis. Meixuezi Tong from the Li lab assisted with *Arabidopsis* transformation. Saemundur Sveinsson analyzed the transcriptomic data.

Chapter 3

I designed and carried out all the analysis. Most of the DNA sequence data was downloaded from GenBank.

Chapter 4

I designed the study with input from Quentin Cronk who performed the crosses and supplied F2 seed. I carried out all the experiments and data analysis. Lina Madilao operated the LC-MS analysis.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIC - Akaike information criterion

AN1 – anthocyanin 1

AN2 – anthocyanin 2

ANS – anthocyanin synthase

AR – anthocyanin rich

bHLH – basic Helix-Loop-Helix

CDS – coding region sequence

CP – Cupani

DFR – dihydroflavonol reductase

DHK – dihydrokaepmferol

DHQ – dihydroquercetin

DHM – dihydromyricetin

F3'H – flavonoid 3'-hydroxylase

F3'5'H – flavonoid 3', 5'-hydroxylase

MC – Mrs Collier

ML – maximum likelihood

PAP1-D – production of anthocyanin pigment 1-dominant

PIC – phylogenetic independent contrast

PL – Painted Lady

ROI – rose intensity

RT – reverse transcriptase

SRS – substrate recognition site

TFs – transcription factors

Acknowledgements

Financial support comes from my supervisor Dr. Quentin Cronk and TAship offered by the botany department. I thank my supervisor for research freedom, guidance and support, and for his understanding during data loss crisis. I also appreciate the input from my committee – Drs. Keith Adam, Jeanette Whitton and Reinhard Jetter.

I am in debt to many individuals at UBC who volunteered their time to help me with this project: Lina Madilao (LC-MS), Maria Diaz Chavez (enzyme assay), Brigitte Cadieux (gene gun), Yoshi Hirakawa (gene gun), Meixuezi Tong (*Arabidopsis* transformation), Junxia Zhang (use of Mesquite). Greenhouse manager David Kaplan helped with maintaining the plants. My labmates Dario Ojeda, Christine Woollacott, Julia Nowak, Saemi Sveinsson, Armando Geraldes and Ji Yong Yang have generously shared their expertise. Saemi analyzed the transcriptomic data. The Otto lab and Doebli lab have kindly shared their resources and equipment.

TA experience at UBC was memorable, thanks to all the instructors, fellow TAs and my students. I am particularly grateful for having Drs. Wayne Goodey, Chin Sun, Carol Pollock and Kathy Nomme as course instructors. In the botany department, I appreciate Ms Veronica Oxtoby for being an angel to all the graduate students! My experience at UBC was made enjoyable by the time spent with roommates and friends. I am also very grateful for people and organizations with similar interests in the outdoor activities, I remember fondly of many hiking trips and field trips in and around BC, many of which led by the Beaty Biodiversity Centre herbarium and Nature Vancouver. I appreciate Olivia Lee and Bill Kinkaid for being both knowledgeable and conscientious during the trips. And lastly, I thank my family and friends overseas for their inspiration and support, to whom I dedicate this thesis to.

Chapter 1. Introduction

1.1 Overview of flower colour shifts in an evolutionary context

Flower colour shifts in this thesis refers to the change in pigment types, caused by genetic or epigenetic mutations, among different individual plants, on an infra-, intra- or inter-specific level. An example of cases that are not included in this concept is when one single plant has different coloured flowers due to ontogenic changes (Weiss, 1995) such as in *Brunfelsia pauciflora* (yesterday-today-tomorrow) and *Hibiscus mutabilis.*

In an evolutionary framework, floral pigmentation plays a role in prezygotic isolation among populations when pollinators preferentially go to one morph and not the other. When associated with other features such as petal cell shape, subtle changes in floral colour may result in a change in pollinator behavior (Comba et al., 2000). This adaptive transition is the major drive, to date, for the study of flower colour shifts in an ecological and evolutionary context (Sobel and Streisfeld, 2013). Plants in which colour shift is associated with a shift in pollinators include but not limited to *Aquilegia* (Whittall et al., 2006, Hodges and Derieg, 2009), *Mimulus* (Bradshaw and Schemske, 2003, Yuan et al., 2013), *Ipomea* (Fineblum and Rausher, 1997, Streisfeld et al., 2011) and *Phlox (Hopkins and Rausher, 2011)*. One converging question is whether repeated evolution of traits is accompanied by parallel changes at biochemical, developmental and genetic levels (Smith and Rausher, 2011).

Floral diversity is often explained by co-evolution between plants and pollinators, where floral variations (such as morphology, colour, patterning and scent) evolved in response to pollinator selection (Kevan and Baker, 1983, Fenster et al., 2004). However, this concept of 'pollination syndrome' is frequently challenged (Waser, 1983, Waser et al., 1996, Johnson, 2006). For example, despite the diverse floral forms in *Gorteria diffusa*, diverging populations share one single pollinator (Ellis and Johnson, 2009). Recently, our understanding is increasing that floral diversity may evolve as an adaptation to non-pollinator-related selecting agents, or through non-adaptive processes such as drift, although evidence in literature supporting the latter is almost non-existent (Rausher, 2008, Armbruster, 2002). Anthocyanins are often associated with tolerance to abiotic stresses (such as heat, drought and high UV exposure), as well as to other biotic stress such as herbivory (Strauss and Whittall, 2006), therefore, adaptive selection via pleiotropic effects is possible. One avenue of future investigation is comparing self-pollinated with cross pollinated plants. Self-fertilized plants may be less prone to pollinator-mediated selection and therefore may respond more to selection from non-

pollination-related agents (Strauss and Whittall, 2006).

Recently there is increasing realization that drift (Fisher and Ford, 1950), or other non-adaptive processes, could also contribute to floral diversity (Rausher, 2008). One problem in demonstrating neutrality in floral colour variation is that lack of evidence for an adaptive case is not evidence for neutrality. It might well be that a significant amount of flower colour variation evolves through genetic drift but that this is under-represented in the literature because researchers fail to report an absence of correlation between pollinator and flower colour variation.

1.2 Biochemistry of flower colour

In angiosperms, flower colour display relies on the pigment production of flavonoids (the entire visible spectrum), carotenoids (yellow to red) or betalains (yellow or red), together with factors such as vacuole pH, copigments and metal ions. The phenolic compounds flavonoids (class phenylpropanoids) are secondary metabolites which are present in both vegetative and reproductive tissues of the plants. Different subclasses of the flavonoids confer a wide spectrum of colour: anthocyanins (orange, red, magenta, violet and blue), aurones and chalcones (yellow), flavones and flavonols (colourless or very pale yellow). In addition to their colouring function, flavonoids are natural anti-oxidant and anti-inflammatory substances which are beneficial for humans (Martínez-Flórez *et al*. 2002). In vegetative tissues, anthocyanins are produced in response to stress since they can absorb light (photoprotection) and scavenge reactive oxygen species (antioxidant) as a protection mechanism (Koes et al., 1994, Mori et al., 2005). The flavonoid pathway is highly conserved in plants and has been studied extensively at both biochemical and molecular levels with most of the pathway genes cloned (Tanaka et al., 2009, Winkel-Shirley, 2001).

UV-A and UV-B irradiation induces free radical, superoxide and H_2O_2 in plants which causes oxidative stress. It is widely accepted that phenylpropanoid compounds are synthesized as an adaptive response to UV exposure, to act as ‘sunscreens’ since they can scavenge free radicals (Caldwell et al., 2007). In particular, the glycosylated form of flavonol and flavone absorb strongly at the UV region. Anthocyanins have two absorption maxima at UV range of 270–290 nm and the visible range of 500–550 nm (Markham, 1982). Red colour are common in juvenile (Woodall and Stewart, 1998), mature (Harborne, 1988) and senescent leaves as anthocyanins can protect expanding leaves from harmful UV radiation. However, it is also argued that anthocyanin accumulation is a by-product of a general

stress response involving other plant hormones, and therefore is non-adaptive (Loreti et al., 2008), although this hypothesis has not been tested.

1.3 Introducing the system: *Lathyrus*

The sweet pea genus *Lathyrus* L. (Fabaceae, Fabeae) includes 150~160 annual and perennial species grouped into 12 or 13 sections (Kenicer et al., 2005). *Lathyrus* is an excellent system to study the molecular basis for colour polymorphism as colour polymorphisms not only exist inter-specifically but also intra-specifically, i.e. within a single species (Pecket, 1960, Broich, 2009). A familiar and representative species of the genus, *L. odoratus* (sweet pea), has a long history of chemical and genetic study of its flower colour that dates back to the early 20th century (Bateson et al., 1906, Bateson, 1913, Punnett, 1923, Fleming, 1925, Beale et al., 1939). Other interests with the sweet pea concern selective breeding for a yellow sweet pea (Hammett et al., 1994, Murray and Hammett, 1998). Breeding systems in *Lathyrus* range from strictly outcrossing in the perennials, facultative autogamy (capable of selfing when pollinators are absent, in most annuals) to cleistogamy (automatic selfing before flower opening, in some annuals) (Brahim et al., 2001). Self-pollination of annuals is usually favored in environments with extremely short growing seasons, where rapid life cycles and time limitation are characteristic (Mazer et al., 2004).

Lathyrus have typical zygomorphic (bilaterally symmetrical) papilionoid flowers which differentiate into an adaxial petal (also known as the standard, banner or vexillum), a pair of abaxial petals (the wings) and a fused pair of lateral petals (the keel). This complicated floral structure allows only specific types of insect to enter and pollinate. In their natural habitat, pollinators of *Lathyrus* species include members of the order Hymenoptera such as *Apis mellifera* (honey bee), *Bombus* and *Megachile* (leafcutter bee), whereas Lepidoptera such as *Hesperia* (a butterfly genus) and thrips cannot operate the flower mechanism and are only visitors to those flowers (Knuth et al., 1906). The sweet pea, *L. odoratus* (**Chapter 2**), is an annual species which is pollinated by a leafcutter bee *Anthidium manicatum* in its natural habitat (Knuth et al., 1906). In *Lathyrus*, whether the flower is pale or dark plays a role in the attractiveness to pollinators and is closely associated with their breeding system (Brahim et al., 2001).

In addition, a novel variety such as the blue rose (Katsumoto et al., 2007) and purple carnation (Fukui et al., 2003) is always sought after in floriculture. Although there are several yellow *Lathyrus* species,

no yellow sweet pea has occurred by spontaneous mutation or through conventional breeding (Hammett et al., 1994). Finding a gene that can eliminate anthocyanin production while maintaining yellow flavonoids or carotenoids could bring in valuable new varieties for floriculture, for instance through genetic engineering.

1.4 Objective of the thesis

The objectives of this thesis are to examine the genetic basis of a red mutant in *L. odoratus* (**Chapter 2**); the polarity of floral colour evolution, in relation to life strategy trait in the genus *Lathyrus* (**Chapter 3**); and the genetic basis for several unpigmented *Lathyrus* taxa that have originated and been maintained under either artificial or natural selection – this is in order to test if genetic parallelism accounts for repeated occurrences of similar traits under different selection processes (**Chapter 4**).

Chapter 2. The Molecular Basis for a Pink Mutant in Sweet Pea

2.1 Introduction

At the beginning of the 20th century, genetic studies of the sweet pea (*Lathyrus odoratus*) inheritance revealed at least 18 hereditary factors, among which the study of flower colour led to the first record of genetic linkage (Punnett, 1923, Bateson et al., 1906, Bateson et al., 1902). Sweet peas are self-compatible and easy to grow which made them attractive and promising to early geneticists. Although the conceptual foundation was not laid during Bateson's time, today the anthocyanin biosynthetic pathway (ABP) and its genes have been well studied (Mol et al., 1998, Winkel-Shirley, 2001, Pecket, 1960) which can be used to reveal the identities of the historically significant hereditary factors.

The sweet pea is native to Sicily where the Mediterranean climate of mild winter and relatively cool summer promotes its persistence. Sweet pea shows a typical architecture of zygomorphic papilionoid flowers, which have differentiated into an adaxial petal (also known as the standard, banner or vexillum), a pair of abaxial petals (the wings) and a fused pair of lateral petals (the keel). The wild type cultivar 'Cupani' ('CP') is named after the Sicilian monk Francis Cupani, who sent the British horticulturist Dr. Uvedale the first batch of sweet pea seeds, which later grew into a worldwide horticultural fascination. Flower *of L. odoratus* is bicolourous with a reddish-purple standard and blueish-purple wings (**Figure 2.1**). Flower colour in sweet pea mainly depends on the type of anthocyanin and flavonol pigments present (Lawrence et al., 1941, Beale et al., 1939). The anthocyanins of the sweet pea are located in the vacuole of the petal epidermis (**Figure 2.1**). The colour of a pink sport 'Painted Lady' ('PL', first appeared in 1731) was found to be controlled by a single locus *A1* (Punnett, 1923), which is tightly linked to a floral shape gene '*A3*', later discovered to be the floral developmental gene *CYCLOIDEA* (Woollacott, 2010). A pink mutant assimilating the cultivar 'Painted Lady' was recorded in the wild by early writers, but not known ever since (Punnett, 1907). It is not known if this pink mutant is ever present in its native habitat Sicily, but if it is, the pink mutation would contribute to the standing genetic variation which may be evolutionarily important (Barrett and Schluter, 2008).

Figure 2.2 shows a partial scheme of the flavonoid pathway. The six major anthocyanidins are classified into three groups based on the number of hydroxyl groups on the B ring: (1) pelargonidin (with one hydroxyl group), (2) cyanidin and peonidin, and (3) delphinidin, petunidin and malvidin. As the number of B-ring hydroxyl group increases, the flower becomes bluer due to a bathochromic shift

whereby the maximal absorption spectra shifts towards longer wavelength (Winkel-Shirley, 2001). Note that blue can also be the effect of co-pigmentation or pH, or the two combined. Through co-pigmentation with metal ions, the pink cyanin (glycoside of cyanidin) can form blue protocyanin in *Centaurea cyanus* (Takeda et al., 2005) and a rise in pH can result in most anthocyanins appearing blue.

The pathway flux to delphinidin production is directed by flavonoid 3', 5'-hydroxylase (F3'5'H, CYP75B, EC 1.14.13.88). Mutations that affect its gene expression is often associated with cyanidin-based colour shift. F3'5'H, along with flavonoid 3'-hydroxylase (F3'H, CYP75A, EC 1.14.13.21) and Flavone synthase II (FNSII, CYP93B), belongs to the cytochrome P450 (CYP450)-dependent monooxygenase family, a large group of membrane-bound enzymes common to plants and animals. Evolutionarily, F3'5'H extends the function of F3'H having additional 5' hydroxylation on the B ring of anthocyanin or flavonol precursors. The C-terminal end of F3'H and F3'5'H distinguishes their functions with substitution of one or two amino acids changing the ratios of 3',5'- to 3'-hydroxylation (Seitz et al., 2007). Most species possess single or two-copy *F3'5'H*, except grape which has as many as 16 copies that might have contributed to the efficient accumulation of anthocyanins in grape skin (Falginella et al., 2010). Rat'kin and Tarasov (2010) reported two copies of *F3'5'H* (*E* and *E1*) in sweet peas, with only one being functional.

In nature, the evolutionary transition from blue to red is asymmetrical with almost no reversal observed across studies (Wessinger and Rausher, 2015). However, reversion to blue may occur when the F3'5'H function is acquired through neofunctionalization of F3'H (Seitz et al., 2006). A colour switch from the blue or purple to pink sometimes involves a pollinator switch, especially when synchronised change in other morphological features are involved. In *Lathyrus* however, there is no definitive case of pollination switch accompanying colour transition from blue to pink/red. But it can be speculated for the bright red *Lathyrus splendens*, which is putatively pollinated by hummingbirds due to its colour whereas its sister species the purple *L. vestitus* is likely to be bee-pollinated. Morphologically however, *L. splendens* is still largely unchanged from other *Lathyrus* (with a landing platform which may still serve bees) and no long corolla tube for typical bird-pollinated flowers has evolved.

In this chapter I aim to identify the *A1* locus of *L. odoratus* var. 'PL' by first profiling the pigment composition and then use candidate gene approach to test if the 'blue gene' *F3'5'H* is associated with

the pink mutant phenotype. For functional verification, I attempted *in vivo* transformation using an *Arabidopsis* mutant line (*PAP1D*) that overexpresses anthocyanin biosynthetic genes as a host. The protocol for transformation in sweet pea is not available whereas transformation into *Arabidopsis* is well-developed and it has no endogenous *F3'5'H*. I did not find any reported study of transforming exogenous *F3'5'H* into an *Arabidopsis* host, so this study will also provide a few considerations concerning the method.

2.2 Materials and methods

2.2.1 Plant materials and crossing

Plants were grown from commercial source and maintained in horticultural greenhouse at UBC with ambient light and temperature. One wild type cultivar 'CP' was crossed with a pink mutant 'PL' to yield six F1 hybrids which were allowed to self-pollinate. 'PL' (the recessive mutant) served as the maternal parents to make it possible to detect self-pollen contamination. A total of 40 F2 plants were scored for colour phenotype and *F3'5'H* genotype.

2.2.2 Pigment analysis

Flowers were harvested just before or at anthesis (S5-6, **Figure 2.3a**) into silica gel. The standard and winged petals were analyzed separately. Each sample contained three flowers from one individual plant that were pooled together as one biological replicate. A total of three biological replicates were included for each cultivar. Water-soluble pigments were extracted with 1 ml methanol: H_2O: acetic acid (85:15:0.5, v/v/v) by soaking overnight in a 4 °C fridge. The solution was centrifuged to retain the supernatant. The extract was acid hydrolysed with 2 M HCl for 30 mins in a waterbath (80 °C) to remove the sugar residues. Precipitates were spinned down and the supernatant was kept in the dark at -20 °C until processing.

HPLC-MS analysis was carried out using a ZORBAX SB-C18 reverse phase column (4.6 × 50 mm, 1.8 μm, Agilent). The crude extract was eluted at a rate of 1 ml min^{-1} with 87% solvent A [H_2O/FA (folic acid), 98:2, v/v] and 13% solvent B (acetyl nitrate/FA, 98:2, v/v) for 4 min, followed by an elution gradient from 23% B for 2 min to 90% B for 1 min, and then back to 13% B. The eluate absorbance at 357—373 nm and 510—530 nm was determined using an Agilent 1100 Diode Array Detector (DAD). Technical quadruplicates with injection volume of 5 μl were performed for all samples. The elute absorbance at 357—373 nm (for flavonol aglycones) and 510—530 nm (for

anthocyanidins) was determined using DAD. Accurate mass of the eluent was then determined by mass spectrometry (MS). Anthocyanidins and flavonol aglycones were examined at the positive and negative ionization modes respectively, following traditions. Peak identification was made by overlaying the MS chromatogram onto the LC chromatogram of a given sample, and pigment was assigned to each sample based on the retention time (T_R) and mass to charge ratio (m/z) of the reference standards and online database (massbank.jp). The criteria for pigment assignment were listed in **Table 2.2.** The relative abundance of an identified pigment was determined by calculating the total area under its peak absorption

2.2.3 Extraction of nucleic acid, cDNA synthesis and PCR

For RNA extraction, fresh petals of *L. odoratus* were harvested at S3-4 late bud stage (**Figure 2.3a**) by flash freezing in liquid nitrogen. Banner and wings were separately analyzed. RNA was extracted with Pure Link Plant RNA Reagent (Invitrogen) following manufacturer's protocol. Total RNA was treated with DNase twice using a Turbo DNA-free kit (Ambion). A first-strand cDNA was synthesized from total RNA (500 ng) using RevertAid First Strand cDNA Synthesis Kit (Fermentas) using supplied oligo$(dT)_{18}$ primer and following the manufacturer's protocol. Genomic DNA from fresh leaf tissues was extracted using a modified CTAB protocol. The *L. odoratus F3'5'H* sequence was obtained by blasting a transcriptome library of *L. odoratus* sequenced and analyzed in our lab (unpublished data). The transcript sequences were then used to design primers. The two ends of *F3'5'H* were confirmed using the First Choice RLM-RACE kit (Ambion) following manufacturer's instructions.

The *L. odoratus* transcriptome library was obtained as follows: total RNA was isolated from flowers of several individuals at the mature bud stage. RNA was treated with DNase twice using a Turbo DNA-free kit (Ambion). The quality of combined RNA from two independent extractions was then determined using a 2100 Bioanalyzer (Agilent). The sample was submitted for whole-transcriptome sequencing by Cofactor genomics (St Louis, Mo.) using the Illumina HiSeq platform. Saemi Sveinsson at the Cronk lab performed the *de novo* assembly using the Trinity assembly programme.

Primers for RT-PCR were designed to span two exons to eliminate false signals given by DNA contamination. *F3'5'H* forward primer F9: 5'-ACCAGGTCCAAAGGGTTATC, reverse primer R8: 5'-AATGTAAGCAGATGGAGACA. Each PCR amplification uses a reaction mixture containing 0.5 mM of each primer, 10 × PCR buffer (Fermentas), 2.0 mM $MgCl_2$, 0.25 mM of each dNTP, and 5 unit

of recombinant Taq polymerase (Fermentas). PCR was performed for 25 or 30 cycles with each cycle condition: 30s at 94°C, 30s at 53°C, 1min at 72°C. PCR products were run on 1.5% agarose gels (0.5×TAE buffer) with 3 µl GelRed (Biotium) for band separation and visualization. PCR product was purified using a QIAquick Gel Extraction Kit following the manufacturer's instructions and sent for Sanger sequencing using an ABI sequencer at the NAPS facility at UBC.

2.2.4 F2 co-segregation analysis

Selfed progenies of the six F1 plants were scored for colour and *F3'5'H* genotypes using the restriction enzyme HhaI (Fermentas), which digests the 332G in 'CP', but not the mutant 332A in 'PL'. Primer used for genotyping were F: 5'-ACCAGGTCCAAAGGGTTATC), R: 5'-CCTTGAAGATCCAACCAAGC. The PCR products of two individuals were sequenced to confirm the genetic background.

2.2.5 Binary vector construction, transformation into *Arabidopsis thaliana*

Two expression constructs, *35S:cpF3'5'H* (wild type) and *35S:plF3'5'H* (mutant type) were generated using the binary vector pCAMBIA-1300. The full-length coding regions of F3'5'H were amplified from 'CP' and 'PL' cDNA, using the forward primer 5'-CCGGAATTCATGATACTCACCCTATACCAAACG and the reverse primer 5'-GCTCTAGACTAAGGAACATAAGCAGATGGAG. The entire CDS of *F3'5'H* were subsequently inserted between the EcoRI and XbaI sites of pCAMBIA-1300. DNA inserts and pCAMBIA-1300 (3 µl, courtesy of Yuti Cheng, Li lab, UBC) were double digested with XbaI and EcoRI at 37 °C for 4 h (DNA inserts) or 12-16 h (vector) to ensure complete digestion. Digested products were recycled from a 1% agarose gel using Gel Extraction Kit (Fermentas). Inserts with appropriate overhangs were then ligated into the binary vector pCAMBIA-1300 with sticky ends. The molar ratio of insert : vector was 3:1 (for 30 ng of 7500 bp vector, 18 ng of 1500 bp insert was added). Ligation was done at 16 °C overnight.

Agrobacterium-mediated gene transformation was carried out with floral dip method (Clough and Bent, 1998). Wild type *Arabidopsis* (*col-0*) was transformed once with 'CP'- and 'PL'-based *35S:F3'5'H* constructs. Basta and Kanamycin resistant plants were selected over two generations for homozygous T2 plants. Two such T2 homozygotes from each cross were then crossed with *PAP1-D* (AT1G56650), a dominant homozygote mutant that overexpresses *PAP1* (MYB75), a positive

regulator of the genes involved in anthocyanin synthesis (Borevitz et al., 2000, Tohge et al., 2005). The cross yielded double heterozygous F1. A control line was generated crossing wild type with *PAP1-D* so that the control plants were heterozygous at *PAP1-D* locus as well. Leaves and roots were then collected from at least six mature F1 plants and control plants into silica gel for pigment analysis. RT-PCR was carried out to verify the expression of the sweet pea *F3'5'H* in the transgenics. Amplification of *Arabidopsis ACTIN1* was used as control with forward primer: 5'-CGATGAAGCTCAATCCAAACGA and reverse primer 5': CAGAGTCGAGCACAATACCG.

2.3 Results

2.3.1 The pink mutant 'PL' lacks delphinidin derivatives, but produces more cyanidin

Just prior to and at anthesis (S5-6 stage, **Figure 2.3a**), both cultivars showed an overall reciprocal change of anthocyanidin and flavonol contents in the two floral parts: anthocyanidins were much higher in the banner than in the wings and the reverse was true for the flavonols, contributing to the bicolourous pattern (**Figure 2.1**). Total anthocyanidin in 'CP' was 1.4~8.7 times that of 'PL', while the total flavonol ratio of 'CP' to 'PL' was 1.3~5.5, indicating more pigment production in 'CP'. The main anthocyanidins in 'CP' were delphinidin and its derivatives in both banner and wings whereas 'PL' produces mainly the cyanidin series, predominately in the banner, with malvidin detected at a low level in the banner (**Figure 2.3b**). The shift from delphinidin to cyanidin indicates a reduced F3'5'H activity in 'PL'. At the same time, the ratio of cyanidin content for 'PL' to 'CP' was 3.4~9.9 in the banner. This indicates the mutation, while abolishing delphinidin pathway, conversely boosts cyanidin production.

Quercetin was the major flavonol in both cultivars, followed by kaempferol. Only a trace amount of myricetin was detected in 'CP' and it was not detected in 'PL'. Note that myricetin is a B ring tri-hydroxylated flavonol (**Figure 2.2**) whose production requires the action of F3'5'H.

2.3.2 A missense mutation in *F3'5'H* is associated with the pink colour in 'PL'

As a lack of delphinidin may be due to a loss of function in *F3'5'H* (**Figure 2.2**), this gene was investigated. *L. odoratus F3'5'H* CDS was determined to be 1548 bp long (513 aa) consisting of two exons (912 and 633 bp in length). The predicted coding sequence of sweet pea *F3'5'H* is 99% identical

to that of *Pisum sativum* and *Medicago truncatula* and 79% identical to *Glycine max*. *Arabidopsis* does not have an *F3'5'H* but it has an *F3'H* (CYP75B1) (which shares 48% identity with the sweet pea *F3'5'H*) encoded by *transparent testa 7* (*tt7*) (Schoenbohm et al., 2000).

No expression difference was found in *F3'5'H* comparing 'CP' to 'PL' using 25- and 30-cycle RT-PCR (**Figure 2.4a**). No difference was detected between different floral parts either. However, by sequencing, it was found that 'PL' carries a missense mutation (G/A) in the first exon 332-bp downstream of the ATG start site, changing the amino acid from glycine to aspartic acid at codon 111, at a highly conserved site in the Substrate Recognition Site1 (SRS1, **Figure 2.5b**) (Seitz et al., 2006, Doyle and Doyle, 1990). F2 co-segregation analysis showed a simple mendelian inheritance pattern for both the 'PL' phenotype and the missense mutation (332 G/A). In a total of 40 F2 plants, the 'PL' phenotype segregated 100% of the time with the point mutation (**Table 2.1,** n=11 mutant plants, and **Figure 2.4b**). Therefore, the G/A missense mutation in *F3'5H* is associated with the 'PL' phenotype, which indicates that it is either the causal mutation or a marker linked to the causal mutation.

2.3.3 Elevated production of cyanidin in the 'PL' *F3'5'H*-transformed *Arabidopsis PAP1-D* line

The *Arabidopsis PAP1-D* (*Production of Anthocyanin Pigment 1-Dominant*, MYB75) is a mutation affecting anthocyanin and other secondary metabolites synthesis (Shi and Xie, 2010). *PAP1-D* plants showed purple colour in all its developmental stages. F1 control plants from a cross between *PAP1-D* and wild type Columbia showed *PAP1-D* phenotype, confirming that *PAP1-D* was a dominant mutant. Contrary to expectation, when the construct expressing *cpF3'5'H* under a *35S* promoter was transformed into a heterozygous *PAP1-D*, the plants failed to produce the blue delphinidin series, but instead produced much higher cyanidin than the control (**Figure 2.6**). The *plF3'5'H*-transformed line produced cyanidins in over 5 fold higher amounts than the control. And this cyanidin production in the *plF3'5'H* transgenic line was also a lot higher than that in the *cpF3'5'H*-transformed line, showing that there is a functional difference between the pl and the cp enzymes. Quercetin and kaempferol accumulation was also a lot higher in the *plF3'5'H* transgenic lines compared to either the control or the *cpF3'5'H* transgenics. All lines produced more kaempferol than quercetin, and no myricetin was detected.

2.4 Discussion

2.4.1 A missense mutation in 'PL' *F3'5'H* occurred at a structurally critical site

The pink colour in 'PL' is associated with a single-base substitution that has led to a change from small non-polar glycine to a polar bulky aspartic acid (G111D). The *A1* locus of 'PL' is therefore identified as *F3'5'H*. The proposal that this is the causal mutation, rather than a marker closely linked to the causal mutation, is supported by further evidence. The same missense mutation (G/A) at the same position of *F3'5'H* exon (332 bp) resulting in a similar amino acid change (glycine to glutamic acid), has been associated with a pink *Pisum sativum* mutant generated by fast neutron bombardment (Moreau et al., 2012). This shows a remarkable parallelism at both phenotypic and molecular levels in two closely related species from the pea family. Although the point mutation in *Lathyrus F3'5'H* does not affect transcription **(Figure 2.4a),** the low amount of malvidin in 'PL' banner suggests it is a leaky, instead of a null, mutation of F3'5'H which limits, but does not abolish, protein function.

The oxidative ability of P450 enzymes depends on electron transfer by the cofactor, which in plants is either a microsomal flavonprotein NADPH-Cyt P450 reductase (CPR, EC 1.6.2.4) or sometimes a haemoprotein cyt b5 reductase (EC 1.6.2.2) (Moller and Lin, 1986, Brugliera et al., 2000). Two electrons are transferred from the flavin coenzymes (FMN and FAD) of the reductase, to oxygen and the cytochrome (Menting et al., 1994b, de Vetten et al., 1999). One of the oxygen atoms is reduced to water while the other is left highly reactive to attack the substrate. Hence, the activity of cyt P450 depends on NADPH, reductase and O_2 (Tanaka, 2006).

Because plant P450 is extremely insoluble when purified, so far no plant P450 has been characterized (Ferrer et al., 2008). However by homology modelling, the structure of four *Arabidopsis* P450 were functionally predicted with reference to the bacteria and mammalian p450, whose crystal structures are resolved (Rupasinghe et al., 2003). Their comparison revealed a high level of structural conservation at the catalytic core even when sequence similarity is as low as 13%: all P450 enzymes share a signature motif F—G-R-C-G in their haem-binding domain (Rupasinghe et al., 2003, Rupasinghe and Schuler, 2006) which, in sweet pea, corresponds to the region **F**GA**G**R**R**I**C**A**G** starting from codon 444 (data not shown).

In addition, SRS1, 2 and 4 (N-terminal) contact the distal aliphatic part whereas SRS4 (C-terminal), 5 and 6 contact the aromatic ring of their respective substrates (Okinaka et al., 2003, Rupasinghe et al.,

2003). In sweet pea cultivar 'PL', the substitution of a glycine (G111) by an aspartic acid in SRS1 might affect substrate binding. Glycine is a unique amino acid as its side chain contains hydrogen instead of carbon, making it more flexible which means glycine can occupy places that are forbidden to all other amino acids such as tight turns in structures. Indeed, by multi-species alignment (**Figure 2.5b**), this G111 corresponds to the *Arabidopsis* CYP75B (F3'H) glycine 105, in **P**NS**G**AKH**M** of SRS1, a residue predicted to contact substrate dihydrokaempferol (DHK) (Rupasinghe et al., 2003). The SRS1 region had a low sequence conservation across the four *Arabidopsis* p450 enzymes and the glycine 105 (in *Arabidopsis*) was not conserved (Rupasinghe et al., 2003) but seems to be specific to the anthocyanin pathway.

F3'H and F3'5'H have broad substrate specificities. In addition to the DHK, substrates that are common to both are naringenin (flavanone), apigenin (flavone) and kaempferol (flavonol). As *Arabidopsis* lacks functional F3'5'H (Forkmann, 1991), we cannot model its F3'5'H directly. But in order to confirm the key residues in F3'5'H, substrate-docking modelling based on *Arabidopsis* F3'H could include more of the common substrates of F3'H and F3'5'H to test if the key residues identified by Rupasinghe et al. (2003), when using DHK as the substrate, are also important when binding other substrates. F3'5'H from sweet pea and *Petunia* could be included in the comparative modelling.

2.4.2 Overexpression of sweet pea wild type *F3'5'H* increases the accumulation of cyanidin in *Arabidopsis*

The *cpF3'5'H* transgenic lines produced more cyanidin than the controls. This is unsurprising as it is well known that F3'5'H may have F3'H activity and that a few amino acid substitutions may convert one to the other (Seitz et al., 2007). The two genes are closely related having diverged from a common ancestor and reversals of function are known in nature (Seitz et al., 2006). *Senecio cruentus F3'5'H* for instance, when tranformed into *Chrysanthemum* fails to produce delphinidin but increases cyanidin production (He et al., 2013). Previous study showed that exogenous F3'5'H failed to accumulate delphinidin in *Chrysanthum*, but instead produced more cyanidin (He et al., 2013).

Heterologous expression of sweet pea F3'5'H was also attempted in both *S. cerevisiae* and *E. coli*, but not enough protein could be obtained for the enzyme assay (data not shown). The main advantage of using yeast is that eukaryotes are similar in the processing and compartmentalization of proteins. But due to codon bias (Hoekema et al., 1987), in some cases plant P450 proteins are difficult to express

unless their sequence is codon optimized (Plotkin and Kudla, 2011, Wessinger and Rausher, 2014). *E. coli* has been used sometimes as an alternative for expression of plant P450. In this case the predicted N-terminal transmembrane helix of the plant P450 is modified to mimic the architecture of the soluble bacteria P450s (Barnes et al., 1991).

I also attempted transient expression of sweet pea F3'5'H *in vivo* by transient expression in epidermal cells, using rupture disk particle bombardment, with GFP as an internal control. But autofluorescence signal (cell death) was stronger than the transformed GFP (data not shown). Other successful transient expression studies in petals based on biolistic methods used a particle inflow helium gun (Vain et al., 1993, Shang et al., 2007, Hellens et al., 2010), which we do not have access to at UBC.

2.4.3 Overexpression of sweet pea *F3'5'H* in *Arabidopsis PAP1D* failed to accumulate delphinidin

PAP1 (MYB75) induces the expression of anthocyanin synthesis genes in all vegetative tissues across multiple developmental stages (Shi and Xie, 2010). Contrary to prediction, *F3'5'H*-transformed *PAP1D* line failed to accumulate delphidin, but did accumulate increased cyanidin (see above). Because *F3'5'H* was successfully expressed in the transgenic lines (**Figure 2.6a**), this could not be due to lack of expression. A lack of enzyme activity is also unlikely because the transformants show increased levels of cyanidin. With the enhanced 3' hydroxylation in the *35S*:*plF3'5'H* transgenic line, it is more likely that the failure to produce delphinidin happened at a functional control level.

The strong 3' hydroxylation coupled with a lack of 5' hydroxylation could have several possible causes, as listed below.

(1) The reaction of F3'5'H could require a NADPH-CYP450 reductase that is absent in *Arabidopsis*. In animals, electron transfer to all the P450s is carried out by a single reductase (Benveniste et al., 1991), and for the 244 P450s in *Arabidopsis* (Bak et al., 2011), there are only two confirmed P450 reductases (ATR): ATR1 and ATR2-2 (Mizutani and Ohta, 1998, Urban et al., 1997, Varadarajan et al., 2010). Therefore, reductases seem to work in a general rather than a specific fashion and hence the *Arabidopsis* reductase should work for the sweet pea F3'5'H. However in *Petunia*, even when another cytochrome reductase is present (Menting et al., 1994a), the inactivation of cytochrome b5 reductase resulted in a reduced F3'5'H activity and reduced accumulation of 5'-substituted anthocyanins, leading to an alteration in flower colour (de Vetten et al., 1999). So it is possible that the sweet pea F3'5'H

might also require a *cyt b5 reductase* that is absent in *Arabidopsis*.

(2) There could be competition of endogenous F3'H with F3'5'H for common substrates. *Arabidopsis* has a functional F3'H encoded by *tt7 (Schoenbohm et al., 2000)* which is active in the *PAP1-D* line which shows purple vegetative tissues. However, the strong apparent boosting of cyanidin production by exogenous F3'5'H indicates that it competes successfully for substrates.

(3) *Arabidopsis* is not an appropriate host for sweet pea *F3'5'H* for reasons other than the lack of a suitable reductase suggested above. Studies have shown that gene origin is important to the successful transformation of *F3'5'H* with some host-donor pairs more successful in producing delphinidin-series than others (Okinaka et al., 2003, Katsumoto et al., 2007, He et al., 2013). One possibility is that substrate specificity for DFR is different between sweet pea and the *Arabidopsis* host.

Although DFR in many species can use all three dihydroflavonols efficiently (Tanaka et al., 1995, Stich et al., 1992), sometimes however, it favors DHK over DHQ and DHM, e.g. in *Antirrhinum* and *Gerbera*, or *vice versa* such as in *Petunia* (Forkmann and Ruhnau, 1987) and *Cymbidium* (Johnson et al., 1999). The proposed mechanism is through a single residue at position 133, with an aspartic acid having a stronger affinity for DHQ and DHM while an asparagine favours DHK (Johnson et al., 2001). *Arabidopsis* DFR is coded by a single copy gene *tt3*, which shares approx. 67% similarity with the sweet pea *DFR* (sequence obtained in Chapter 4) on a protein level. Aligning *tt3* with the sweet pea *DFR* I found that sweet pea *DFR* indeed has an aspartic acid (DHQ and DHM inclined) whereas *tt3* has asparagine (DHK inclined) at position 133 (data not shown). The implication is that being able to act on DHQ and DHM directs the pathway to cyanidin and delphinidin production, respectively (Figure 2), whereas DHK is in the pathway to pelargonidin (but is also precursor to cyanidin). However, even though the *Arabidopsis* DFR is not apparently optimized for converting DHQ (cyanidin pathway), cyanidin is still produced in large amount in the *PAP1-D* mutant lines, indicating that substrate specificity has a rather minor effect. Furthermore this does not explain the discrepancy in the *cpF3'5'H* transgenics between cyanidin (from DHQ) production and a lack of delphinidin (from DHM) production.

2.4.4 Alteration of the SRS1 domain of sweet pea *F3'5'H* greatly increased the accumulation of cyanidin in a heterologous system

The cyanidin content in the mutant *plF3'5'H* transgenic lines was not only 5- to 6-fold that of the control, it was also considerably higher than the wildtype *cpF'3'5'H* transgenic lines. This indicates,

at least, a functional difference between the *cpF3'5'H* and *plF3'5'H* with the latter probably being optimized for the 3' hydroxylation. All F3'5'H enzymes have some F3'H activity (Seitz et al., 2006). It is remarkable that the alteration of a single amino acid in the F3'5'H greatly increases cyanidin production over wild-type, indicating that this mutation is physiologically significant. In sweet pea this mutation apparently compromises delphinidin production, but it appears to make cyanidin production more efficient, i.e. the mutation appears to turn the enzyme into a better 3'hydroxylase while eliminating 3'5'hydoxylase activity.

As previously mentioned the mutant amino acid change from the highly conserved glycine (G) to aspartic acid (D) in SRS1 is structurally homologous to the key glycine residue in *Arabidopsis* F3'H predicted to contact substrate DHK (Rupasinghe et al., 2003). DHK is a precursor for cyanidin (and delphinidin, **Figure 2.2**). By altering this glycine in sweet pea F3'5'H to aspartic acid, it is possible that the enzyme might be toggled from primary F3'5'H activity (and minor F3'H activity) to a relatively efficient F3'H (with little F3'5'H activity), through unknown enzyme kinetics. The resulting change is a 'gain-of-function' to make more cyanidin while 'loss-of-function' in terms of delphinidin. This result would follow if aspartic acid is a better substrate binding residue for DHK (and/or DHQ) than glycine at this position. This would also predict that PL would have more cyanidin in its petals than 'CP', as the mutant F3'5'H is predicted to act as an F3'H. This is exactly what is observed (**Figure 2.3b**).

The shift in F3'5'H activity towards F3'H might be accentuated by a release of the cyanidin pathway to competition from the delphinidin pathway. DHQ (precursor of cyanidin) and DHM (precursor of delphinidin as well as cyanidin) may compete for *cpF3'5'H* which can bind to both whereas *plF3'5'H* apparently has a compromised 5' hydroxylation ability so might only efficiently bind to DHK/DHQ. In addition, in order for the transgenics to produce delphinidin and its derivatives, it might be necessary to optimize the environment by down-regulating competing pathways (Tanaka and Ohmiya, 2008).

Since a single transformation event only incooperates a single copy of the Ti plasmid into the *Arabidopsis* genome, the increased pigment accumulation in both transgenics could also be due to the following reasons: **(1)** The insertion point was close to the upstream of a positive regulator (or its promoter/enhancer region) of the anthocyanin pathway, such as any gene of the TTG1/bHLH/MYB complex (Gonzalez et al., 2008), which increased anthocyanin accumulation (**Figure 2.6b**). However,

the two transgenics (*cpF3'5'H* and *plF3'5'H*) served as two replicates testing against such positional effect. **(2)** The insertion disrupted a negative regulator of the flavonoid pathway, such as *SPL9* (Gou et al., 2011) – the only known negative regulator so far. However, since *SPL9* regulates anthocyanin and flavonol production reciprocally (Gou et al., 2011), an increase in anthocyanin accumulation would be accompanied by a reduced flavonol production, which was not the case (**Figure 2.6b**). **(3)** Unknown positional effect for the *plF3'5'H* transgenics explains the overall increase in both anthocyanin and flavonol production. It is well-known that the metabolic flux in the flavonoid pathway is governed by substrate competition between FLS and DFR, it is a tug of war to produce either flavonol or anthocyanin (Davies et al., 2003). In response to stress, *Arabidopsis* increased its anthocyanin, rather than flavonol production (Kovinich et al., 2014). Therefore, only when there is an increased supply of dihydroflavonols that both types of pigments could then increase abundantly. But this cannot happen within this closed system without external feeding of the substrates. Furthermore, since in *Arabidopsis*, the early and late stages of the flavonoid pathway is separately regulated by different TFs (Stracke et al., 2007, Martin and Gerats, 1993), it is highly unlikely that a single insertion would enhance the genes involved in both early (flavonol production) and late (cyanidin production) stages leading to what is observed in the *plF3'5'H* transgenics (**Figure 2.6b**). Therefore, if positional effect was to explain the enhanced flavonol production in *plF3'5'H*, the increased cyanidin accumulation is unlikely due to this same effect. In future, more independent transformations will confirm or disprove the positional effect.

2.4.5 Evolutionary implications

A missense mutation in *F3'5'H* caused the transition from purple to pink in a sweet pea cultivar studied a century ago (Bateson et al., 1906, Bateson et al., 1902). Colour change has often been associated with point mutations, such as in the white Kermode bear population inhabiting the British Columbia pacific rainforest (Ritland et al., 2001) and in yellow-rooted Cassava (Welsch et al., 2010). In nature, rare mutations of large effect and infinitesimal mutations with small effects are both needed to achieve complex adaptations (Mayr and Provine, 1980). However, in the case of flower colour transition, the system is so efficient that due to the clearly distinguishable pigments, often only a few mutations at a few genes are needed to give an obviously discernible effect.

The literature is abundant with the association between *F3'5'H* and the evolutionary trajectory from

blue to pink/red. Mutation in *F'3'5H* is common because it lacks associated pleiotropic effects. This may lead to genetic constraints on further evolutionary trajectories, making reversal rare for the blue to red transitions (Wessinger and Rausher, 2015). The inactivation of the delphinidin branch in *Chrysanthemum*, *Dianthus*, *Rosa*, *Ipomoe*a are all caused by the absence of F3'5'H activity (Katsumoto et al., 2007, Fukui et al., 2003). And in *Penstemon*, parallel evolution from blue to red is accompanied by repeated degeneration of the *F3'5'H* locus (Wessinger and Rausher, 2015). When *F3'5'H* is associated with the blue to red transition, the mutation often happens at the *cis* region of the gene, being either a functional mutation in the coding region in *Gentiana* (Nakatsuka et al., 2006) and *Penstemon* (Wessinger and Rausher, 2015), or in the *cis* regulatory element in *Saintpaulia* (Sato et al., 2011) and *Phlox* (Hopkins and Rausher, 2011).

However, in soybean (*Glycine max*), a mutation in *flavanone-3-hydroxylase (F3H)* caused a blue to pink transition whereas the degenerate *F3'5'H* is associated with a white mutant (Zabala and Vodkin, 2007). And in *Glycine soja*, an amino acid change in the *F3'5'H* coding region, decreasing the enzyme activity, only led to a change in shade from purple to light purple (Takahashi et al., 2010). On the other hand, not all blue flowers are delphinidin based: The blue colour in Morning glory (*Ipomoea*) is from cyanidin and the transition from blue to red in the genus was due to *F3'H* mutations which led to pelargonidin production (Hoshino et al., 2003). So we can see some degree of predictability but a lack of certainty for the *F3'5'H*-blue association in literature. Blue to red transition in *Lathyrus* occurs both at and below the species level and further study in the system will contribute to the emerging generalisation while elucidating any ecological and evolutionary significance. For instance, the red *L. splendens* is likely hummingbird-pollinated, *L. nevadensis* has both blue and pink morphs. Is the same gene (*F3'5'H*) responsible for colour transition under different selection pressure and evolutionary contexts?

Table 2.1 F2 co-segregation analyzed with respect to the colour phenotype and F3'5'H genotype in a total of 40 F2 individuals obtained from the cross 'CP' (purple) × 'PL' (pink). Pink is the recessive phenotype (purple: pink = 3:1) and a ('PL' F3'5'H marker) is the recessive genotype (AA:Aa:aa 1:2:1, A is the 'CP' F3'5'H allele).

Individual	Genotype	Phenotype	Individual	Genotype	Phenotype
A1	**aa**	**Pink**	A25	**aa**	**Pink**
A2	**aa**	**Pink**	A26	Aa	Purple
A3	Aa	Purple	A27	Aa	Purple
A4	Aa	Purple	A28	Aa	Purple
A5	**aa**	**Pink**	A29	AA	Purple
A6	Aa	Purple	A30	**aa**	**Pink**
A7	**aa**	**Pink**	A31	AA	Purple
A8	Aa	Purple	A32	**aa**	**Pink**
A9	**aa**	**Pink**	A33	Aa	Purple
A10	**aa**	**Pink**	A34	Aa	Purple
A11	Aa	Purple	A35	Aa	Purple
A12	Aa	Purple	A36	Aa	Purple
A13	Aa	Purple	A37	AA	Purple
A14	Aa	Purple	A38	AA	Purple
A15	Aa	Purple	A39	Aa	Purple
A16	Aa	Purple	A40	**aa**	**Pink**
A17	Aa	Purple	A41	**aa**	**Pink**
A19	Aa	Purple	A42	Aa	Purple
A20	AA	Purple	A43	AA	Purple
A21	AA	Purple	A44	Aa	Purple

Table 2.2 Peak assignment for the analysis of the aqueous methanol extract of the flowers of *L. odoratus*, based on the flavonoid aglycone standards. T_R (retention time) decreases as polarity of the molecule gets higher. More polar molecules elute faster in a reverse phase LC. m/z (mass to charge ratio) calculation is based on the mass database at massbank.jp. +/- indicates the mode the compound is detected with. Absorption maxima followed manufacturer's certificate of analysis.

Pigment	t_R (min)	m/z	UV λ_{max} (nm)
Cyanidin	1.4	287(+)	538
Peonidin	2.8	301(+)	537
Delphinidin	0.9	303(+)	548
Malvidin	3.1	331(+)	547
Petunidin	1.8	317(+)	547
Quercetin	4.5	301(-)	372
Kaempferol	5.8	285(-)	365
Myricetin	2.8	317(-)	374

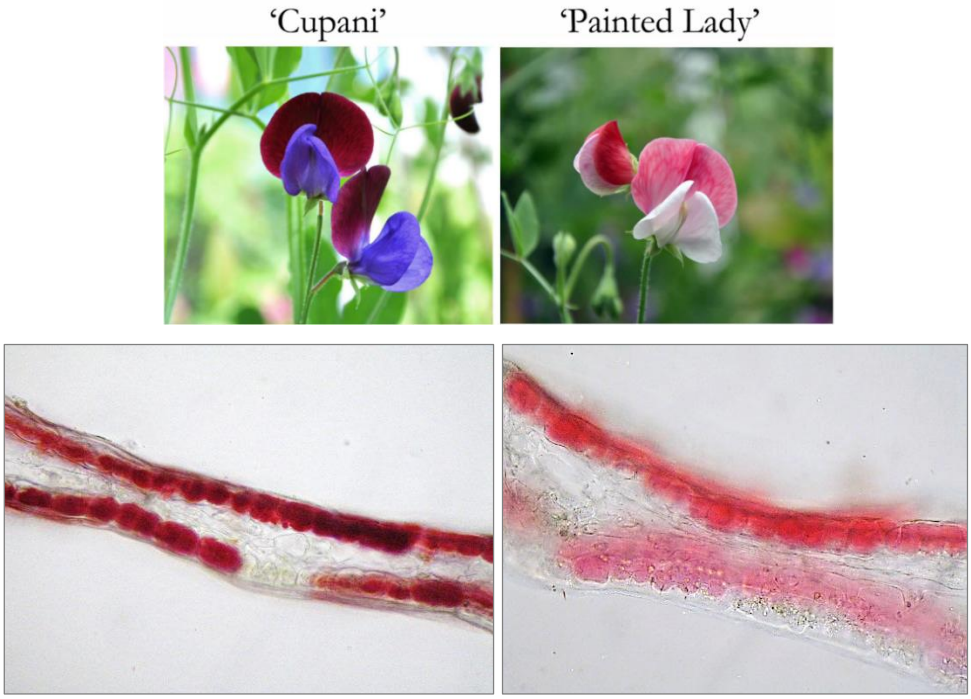

Figure 2.1 Top panel: *Lathyrus odoratus* wild type cultivar 'CP' and pink mutant 'PL'. Bottom panel: transverse section of the banner petals showing an epidermal position of the anthocyanins in 'CP' (left) and 'PL' (right). Magnification: 200 ×.

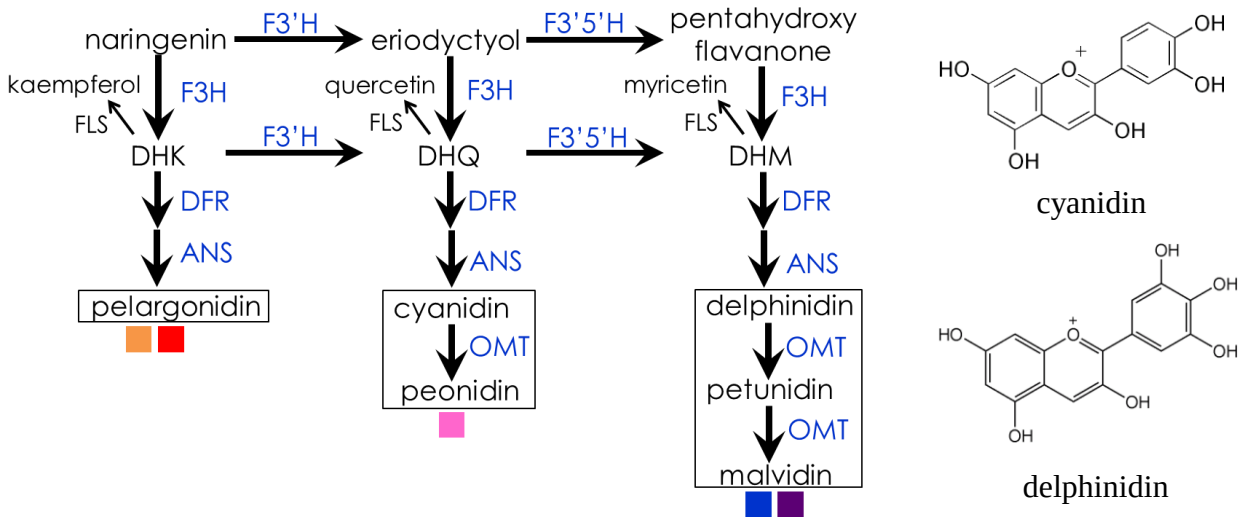

Figure 2.2 Intermediates in the flavonoid pathway starting with flavanones and leading to anthocyanidin synthesis. Flavanones (naringenin, eriodyctyol and pentahydroxy flavanone) are converted by flavone 3-hydroxylase (F3H) into dihydroflavonols DHK (dihydrokaepmferol), DHQ (dihydroquercetin) and DHM (dihydromyricetin), which can be further converted to either the flavonols by flavonol synthase (FLS), or to anthocyanidins, catalysed by dihydroflavonol reductase (DFR) and anthocyanin synthase (ANS). Further methylation of cyanidin and delphinidin into their respective derivatives are carried out by *O*-methyltransferase (OMT). The colour of the end products of each pathway is indicated, anthocyanidins (indicated in brackets) and flavonols (kaempferol, quercetin and myricetin) differ in the number and type of substitution in the B ring. Kaempferol and pelargonidin are monohydroxylated at the C3 position, quercetin and cyanidin are dihydroxylated at C3 and C4 positions, myricetin and delphinidin are trihydroxylated at C3, C4 and C5 positions. Flavonoid 3'-hydroxylase (F3'H) directs the pathway flux to cyanidin whereas flavonoid 3'5'-hydroxylase (F3'5'H) catalyses the dihydroxylation and directs the pathway to producing the delphinidin series. The hydroxylation patterns on the B ring of cyanidin and delphinidin, the key chemicals in this study, are shown on the right (source from Wikipedia).

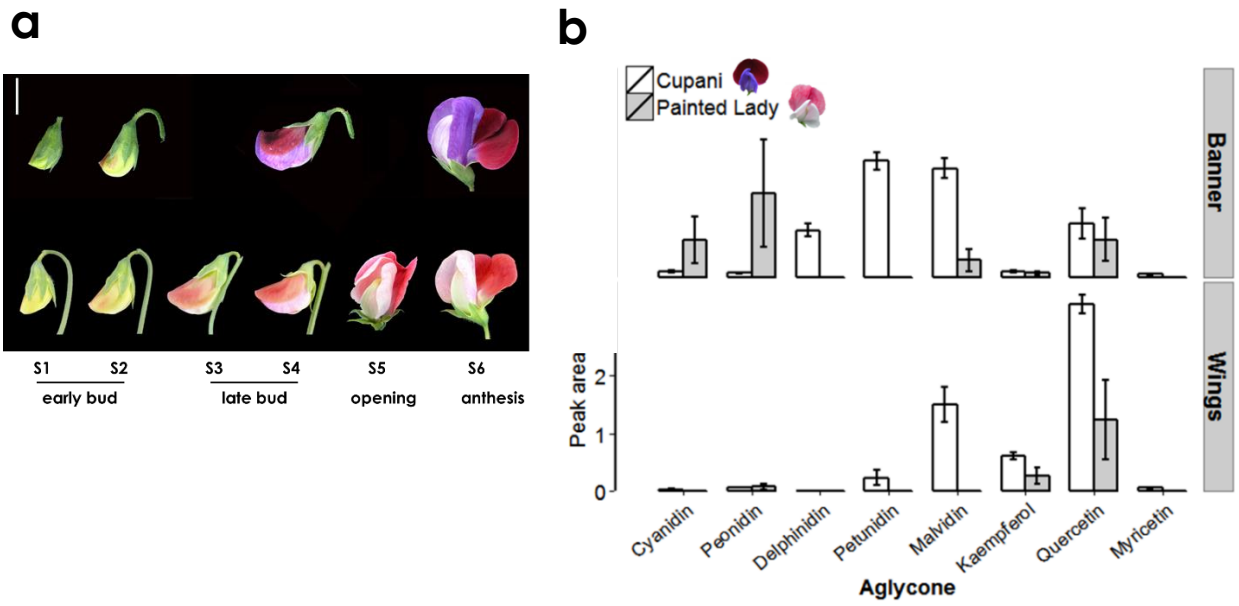

Figure 2.3 (a) The six flower ontogeny stages from bud initiation (S1) to anthesis (S6). Top panel, *L. odoratus* cv. 'CP' (wild type). Bottom panel, *L. odoratus* cv. 'PL' (pink mutant). S1, small bud without visible anthocyanin, bud curving down towards the pedicle. S2, anthocyanin pigment starts to form. S3, anthocyanin is clearly visible. S4, a mature bud forms and it continues to curve away from the pedicle as it enlarges until finally at S5, opening of the bud accompanied by an upright positioning of the flower now along the same axis as the pedicle. S6, full opening of the flower with clear separation of banner and wings. Scale bar represents 1 cm. **(b) Pigment profile for 'CP' and 'PL', measured at S5-6 stage.** Relative abundance is measured as peak area under each identified MS peak.

Figure 2.4 (a) RT-PCR of *F3'5'H* expression in wild type 'CP' and pink mutant 'PL' showing no substantial difference. b, banner petal, W, wings. **(b) F2 co-segregation analysis digestion gel.** The wild type allele will be digested into two fragments by restriction enzyme HhaI whereas the mutant allele is not cleaved at all. 'CP' phenotypes were a mixture of 22 heterozygotes and 7 homozygotes for the *F3'5'H* 'CP' allele (not all 'CP' individuals were shown here).

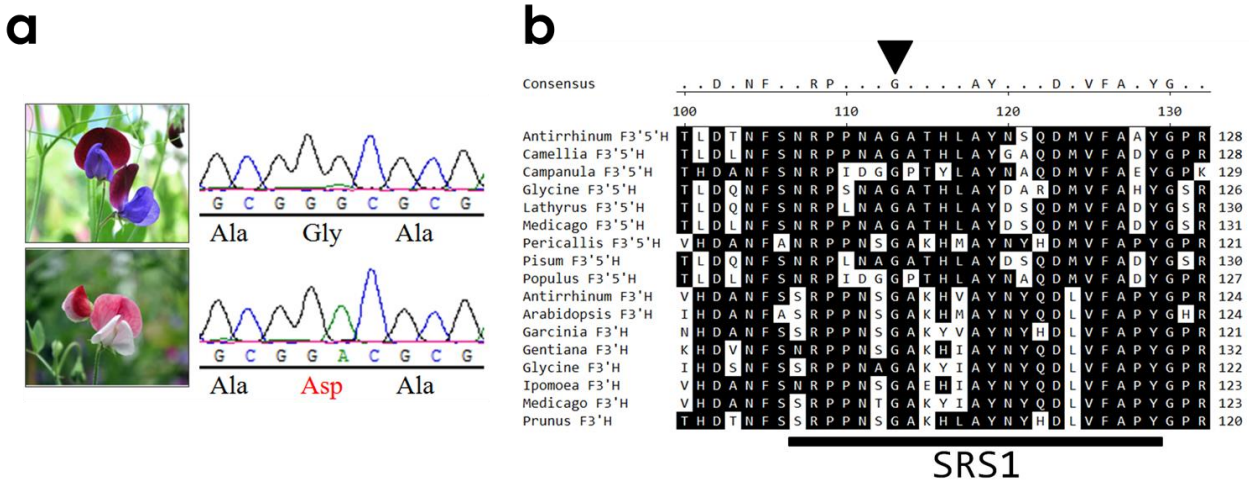

Figure 2.5 (a) Chromatogram of *F3'5'H* showing the missense mutation changing the small uncharged amino acid glycine (in 'CP') into a large and negatively charged aspartic acid (in 'PL'). (b) Multi-species alignment showing the single nucleotide substitution is in a highly conserved position of the SRS1 region across plant species.

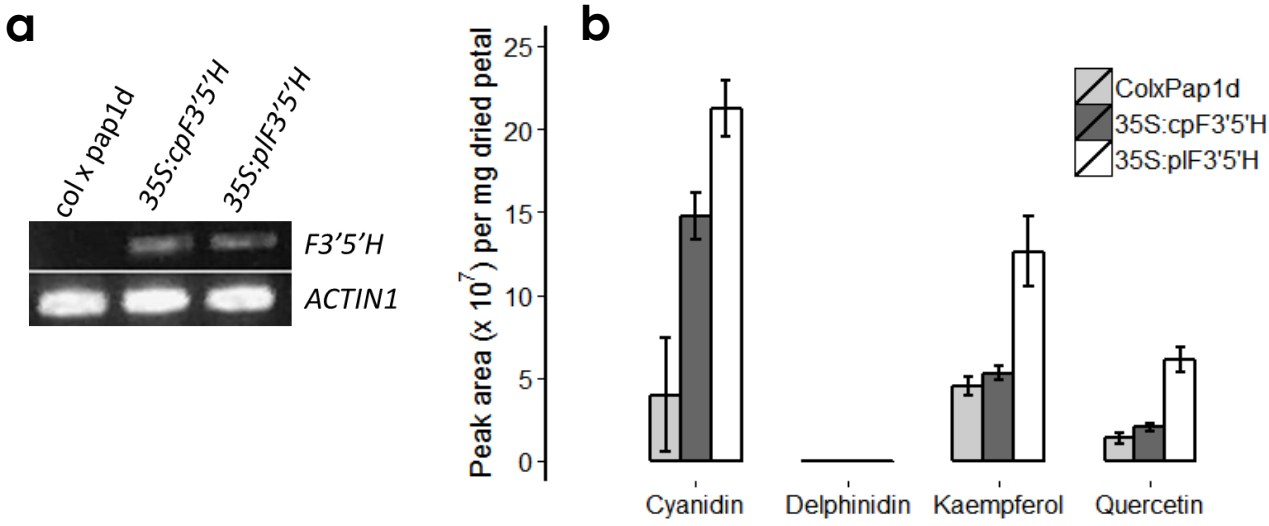

Figure 2.6 *Arabidopsis PAP1D* (a dominant mutant line that produces anthocyanins constitutively) transformed with sweet pea *F3'5'H* showed elevated levels of cyanidin production, but no delphidin production. (a) RT-PCR showing the expression of sweet pea *F3'5'H* were successfully induced in the transgenic lines. (b) Abundance of flavonoid aglycones in leaves and stems from the control plants and the transgenic lines. Col × pap1d: control, from crossing Columbia and homozygote *PAP1-D*. 35s:cpF3'5'H, *PAP1D* transformed with wild type *F3'5'H*. 35s:plF3'5'H, *PAP1D* transformed with mutant *F3'5'H*.

Chapter 3. Phylogenetic Reconstruction for Floral Colour, Life History Trait and Pattern in *Lathyrus*

3.1 Introduction

In this study, I exploited the system of *Lathyrus* L. (Fabaceae, tribe Fabeae Rchb.), a large genus containing ca. 150~160 annual and perennial species grouped into 13 (Kupicha, 1983), and recently 11 sections (Asmussen and Liston, 1998, Kenicer et al., 2005, Shehadeh, 2011, Lewis et al., 2005), to survey colour and pattern, putative visual cues for pollinators. Colour shifts in the primarily insect-pollinated *Lathyrus* are common within and between species, an ideal system for genetic comparisons of parallel and convergent phenotypic evolution (Conte et al., 2012, Wessinger and Rausher, 2012). I will investigate the polarity of colour evolution in *Lathyrus*, and broadly visualise colour, pattern and life strategy shifts in this system, identify 'replicate' cases of phenotypic evolution (**Chapter 4**). 'Pattern' refers to the petal colouration of bicolour (different colour in banner and wing petals) versus concolour (same colour throughout). The banner petal is larger in size and serves as main floral display while the wing petals serve as landing platforms. Therefore contrasting visual cues in the two types of petal likely serve as visual cues to pollinators.

The most common polymorphisms are pink, blue and white, as in many wild flowers (Griffiths and Ganders, 1983). The genetics of red shift was investigated in **Chapter 2** and that of the 'white shifts' will be investigated in **Chapter 4**. Some *Lathyrus* species also possess UV patterns that add to colour diversity (personal observation). The genetics of *Lathyrus* colour dates back to the beginning of modern genetics with Bateson, Saunders and Punnett (Bateson et al., 1906, Punnett, 1907, Bateson, 1913, Punnett, 1923). However, there is a lack of effort putting their discovery in a modern evolutionary framework, such as to investigate the frequency and reversibility of colour shifts, which colour is ancestral, which gene(s) is involved in colour transitions, and above all if any colour transition is matched with pollinator shift. With the full characterization of the anthocyanin pathway, there is now the potential to rediscover *Lathyrus* genetics in a modern context.

Annual *Lathyrus* are autogamous, capable of self-fertilization when pollinators are absent whereas perennials are strictly outcrossing (Brahim et al., 2001). However autogamous annuals show higher fertility under outcrossing conditions compared to selfing conditions, as measured by pods per flower and seeds per pods (Brahim et al., 2001), indicating a facultative autogamous system in annual *Lathyrus*. In their study, even *L. odoratus* (which is consistently self-pollinating in cultivation, when

pollinators are absent) set significantly more seeds from outcrossing in a natural habitat.

The highest diversity of the genus lies in the Eastern Mediterranean region where the tribe Fabeae originated (Schaefer et al., 2012, Kenicer et al., 2005). Many members of the genus are of economic and agronomic interest as ornamentals (*L. odoratus, L. latifolius*), forage (*L. ochrus, L. articulate*s) and human food (*L. sativus*). Some *Lathyrus* can persist in arid conditions, making them ideal crops in ecologically vulnerable regions especially in the phase of increasing frequency of extreme climate. Recent biogeographic evidence has pointed to an Eastern Mediterranean origin for *Lathyrus*, which subsequently reached the new world either via long distance dispersal across the Atlantic or dispersal from Asia over the Bering land bridge (Kenicer et al., 2005, Schaefer et al., 2012). More recently some species may have been transported across the Bering Strait by the Inuit who had traditional uses such as the seeds (Kenicer et al., 2005, Schaefer et al., 2012). However without reliable fossils of Fabeae it is hard to infer precise colonization events (Lavin et al., 2005). Previous phylogenetic reconstructions using maximum parsimony (Kenicer et al., 2005) and maximum likelihood models based on concatenated nuclear and chroloplast data (Schaefer et al., 2012) have improved the traditional taxonomy set by Kupicha (1983). Here a phylogeny for *Lathyrus* based on a multispecies coalescent model was constructed and was used for ancestral state reconstruction (Pérez et al., 2006).

3.2 Materials and methods

3.2.1 Data collection and taxon sampling

I downloaded GenBank data from contributions of Kenicer et al. (2005) and Schaefer et al. (2012) plus the Mediterranean species *L. chloranthus* and *L. chrysanthus*. Molecular data included five plastid genes (the protein coding *mat*K, *psb*A and *rbc*L genes, the *trn*L intron plus flanking spacer *trn*L-F and the *trn*S-G intergenic spacer), and one nuclear locus (the nuclear ribosomal internal transcribed spacer (ITS) plus the 5.8S coding regions of nuclear ribosomal DNA). The five chloroplast loci were concatenated into a single locus for analysis. Sequences for 105 taxa were either downloaded from GenBank or amplified using genomic DNA (**Table 3.1**). Total genomic DNA was extracted using a modified CTAB method with an addition of 2% β-mercaptoethanol to the extraction. PCR amplifications were purified using PCR Purification kit (Fermentas). The primers and protocols for amplification, purification and sequencing of PCR products followed Kenicer et al. (2005). Sequences were then validated using BLAST to compare with GenBank. Samples were selected so that every

taxon is present in both combined plastid and nuclear ITS datasets. A total of 105 samples across 101 effective species (92 *Lathyrus*, two *Pisum* and 11 outgroups) were included in phylogenetic estimation.

3.2.2 Phylogenetic reconstruction of *Lathyrus*

Sequences were aligned using MUSCLE 3.8 (Edgar, 2004) and manually adjusted in Mesquite 2.75 (Maddison and Maddison, 2011). Poorly aligned regions were excluded. Gaps were treated as missing data. The final matrices included two loci: 641 aligned positions of the nuclear ITS region (163 informative sites) and a combined plastid region of 4572 characters (536 informative sites). A species tree was generated with Bayesian multi-species coalescent approach implemented in BEAST v1.8.1 (Heled and Drummond, 2010). This method assumes that incomplete lineage sorting is the only source for gene-tree incongruence. Both loci were assigned GTR + G substitution models, as determined using jModeltest 2.0 (Darriba et al., 2012, Guindon and Gascuel, 2003). A lognormal relaxed clock rate of 1.0 was assigned to the plastid locus while the nuclear locus was estimated relative to it. A birth-death process was chosen as prior for the speciation process. The plastid ploidy type was set to 'mitochondrial' as plant chloroplasts and animal mitochondria are both maternally inherited haploid organelles. The nuclear ploidy type was set to 'autosomal nuclear'. All other priors were left to default settings (Drummond and Rambaut, 2007). Two sets of Markov Chain Monte Carlo (MCMC) were run simultaneously using Metropolis coupling, each with one cold and three heated chains. Two independent replicates were performed. Each run included 800,000,000 generations with trees sampled every 10,000 generations. The first 25% trees from each replicate were discarded as burn-in and convergence was checked by examining trace and ESS higher than 200. The remaining trees were combined from the two replicates and used to build a maximum clade credibility tree accompanied by posterior probability (PP).

3.2.3 Ancestral reconstruction of floral traits and correlation analysis

Colour and pattern assignments followed multiple sources, cross checked with each other. Major sources include: the manuscript treatment of *Lathyrus* in Flora of North America (Broich, unpublished work) , Flora Iranica (Chrtkova-Zertova et al., 1979), online databases including Encyclopedia of Life (eol.org), JSTOR Plant Science (plants.jstor.org), Flora of China and Pakistan (efloras.org, eflora.cn), Calphotos (calphotos.berkeley.edu), SEINet (swbiodiversity.org) and field pictures. Evolutionarily, loss of strong anthocyanin production results in a colour shift from pink, red, blue or purple to white

or yellow. A binary character state for floral colour was therefore assigned as 1 (anthocyanin colouration present: pink to purple) or 0 (anthocyanin colouration largely absent: white to yellow), i.e. a white flower with pigmented veins (or nectar guides) is coded as 0. Many *Lathyrus* have bicolourous corollas. To be consistent, only the major floral colour indicating maximum anthocyanin-producing ability was used. Life cycle was assigned as 1 (perennial) or 0 (annual). Floral colour pattern was assigned as 1 (concolourous) or 0 (bicolourous). Floral colour, life cycle and pattern traits as three binary characters were each reconstructed using evolutionary models of maximum likelihood (ML) implemented in Mesquite 2.75 (Maddison and Maddison, 2011). Analyses were carried out on the highest likelihood tree with GTR + G branch lengths. I reconstructed all three traits with modeled character evolution assuming asymmetric probabilities of gains and losses (MK2) as determined by log likelihood ratio test between MK1 (symmetric evolution) and MK2 models.

Pagel's lambda (λ) (Pagel, 1999) was estimated as a strength for traits' phylogenetic signals (the tendency for closely related species to share similar traits). Outgroups were excluded for the calculation of Pagel's lambda. Maximum likelihood λ was fitted using *fitDiscrete* command in R (v 3.2.2) package *geiger* (v 2.0.6) (Harmon et al., 2008). Independence of the two binary characters in colour and life cycle was tested using Pagel's 1994 correlation analysis performed in Mesquite with 500 simulations (Maddison, 1990). But since species may not be phylogenetically independent, non-independent data points were corrected by Phylogenetic Independent Contrast (PIC) using the *pic* command in *ape* (v 3.3) (Paradis et al., 2004). Correlation coefficient corrected for PIC was then calculated using the *glm* command in *nlme (Pinheiro et al., 2013).*

3.2.4 UV photography

Hymenopterans view the world primarily in lower wavelength range, with a conserved trichromatic vision peaks at 340 (UV), 430 (blue) and 535 (green) nm (Briscoe and Chittka, 2001). The UV reflectance pattern in Figure 3 were shot with a digital camera (Fuji finepix J10) with non-coated lens to let both IR and UV pass through), fitted with two lens that bypasses UV and infrared light. UV lens (X-Nite 330) has a peak transmission at 330 nm and UV+IR (X-BP1) passes 330-630 nm 930-1400 nm. The photos were taken in a dark room where the only source of light is a UV torch emitting single wavelength light at 375 nm.

3.3 Results

3.3.1 Main phylogenetic results

Multispecies coalescent analysis gave the tree shown in **Figure 3.1**. Consistent with previous studies (Kenicer et al., 2005, Schaefer et al., 2012, Shehadeh, 2011), *Lathyrus* and *Pisum* form a monophyletic group and the outgroup *Vicia* is paraphyletic. The deepest divergence, albeit weakly supported, is between a clade containing the sect. *Clymenum* except *L. gloeospermus,* plus the monotypic sect. *Neurolobus* (an autogamous perennial), plus *Pisum*, and the rest of the *Pisum-Lathyrus* clade. This early divergent group, and the positions of taxa within it, is the same as in Schaefer et al.'s (2012) concatenated ML tree. *L. gloeospermus* is separated from the rest of the sect. *Clymenum*, a placement that has always been problematic (Asmussen and Liston, 1998). *L. gloeospermus* and *L. nissolia* was placed as sister group in Schaefer et al., with a low BS support. In this study they are placed as successive outgroups to the larger clade, with low support. The next level of divergence is between members of the sect. *Lathyrus* (incluyding *L. angulatus* in both analyses), and the clade including sect. *Lathyrostylis*, a pattern again consistent with Schaefer et al. study. Furthermore, sect. *Orobus* in both studies were revealed as a recently diverging section. This is significant to the present study as species of the series *Lutei*, with many autapomorphies, including yellow racemes, belong to sect. *Orobus*, a section widely distributed in the Old and New Worlds.

The biggest difference between this analysis and Schaefer et al.'s topology lies in the relationship among the recent-diverging sections: in their study, sect. *Orobus* is sister to a clade comprised of sect. *Pratensis* and sect. *Aphaca*, and this clade forms a polytomy with sect. *Notolathyrus* and sect. *Linearicarpus*. However, in this study, sect. *Orobus* is sister to the sect. *Notolathyrus* – sect. *Pratensis* clade (low support), and together this clade is sister to *Linearicarpus.* Considering the low support in both analyses, the difference should be interpreted as a "soft disagreement" until more nuclear data can be included. At present the Schaeffer et al. study, as it is more densely sampled, should be considered the most reliable, given the importance of sampling in accurate phylogenetic reconstruction (Heath et al., 2008). Nonetheless, the four major sections *Orobus, Pratensis, Notolathyrus, Linearicarpus* were placed together in both studies.

It has been suggested from traditional taxonomy, that sect *Lathyrostylis* should be included in sect. *Orobus*. However, consistent with Schaefer et al. (2012), *L. pannonicus* was the only species from *Lathyrostylis* that was placed in *Orobus*. Therefore, like Schaefer et al., this study challenges

traditional taxonomy by providing evidence that *L. tukhtensis* and *L. sphaericus* should be placed in sect. *Lathyrostylis*. Not included in Schaefer et al.'s study, two yellow East Mediterranean species in the sect. *Lathyrus* - *L. chloranthus* and *L. chrysanthus* (Sahin et al., 2000, Ayaz and Ertekin, 2008) were grouped together as sister taxa in the BEAST topology, as expected from traditional taxonomy (**Figure 3.1**).

In summary, the ML concatenated tree constructed by Schaefer et al. (2012) and BEAST tree constructed here agreed on the placement of four early diverging sections and the coarse segregation of major sections. We also agreed on the detailed placement of some problematic/ambiguous species. Relationships within each section are less important to ancestral reconstruction to follow. However, most of the conflicting branches received low support in both studies. It seems that the similar genetic data used by the two methods are robust to differences in phylogenetic models. As there are certain taxa important for this study that are not included in the Schaefer et al. study, the ancestral reconstruction in this chapter will use the BEAST tree.

3.3.2 Ancestral reconstruction and correlation analysis

For ancestral reconstruction alone, I included nine white, 13 yellow and 70 anthocyanin-rich (AR: pink to blue and purple) *Lathyrus* and *Pisum*. The annual, AR, concolourous form is the ancestral state, inferred with both ML and MP constructions (MP results not shown). ML reconstructions gave probabilities of over 80% for the colour trait, 100% for the life cycle trait and over 70% for the pattern trait (**Figure 3.2 and 3.3**). Under parsimony ancestral state reconstruction, loss of colour happened at least 16 times with no reversal (**Figure 3.2**). Transition from annual to perennial happened at least four times with three reversals (**Figure 3.2**). For pattern reconstruction, transition from concolourous to bicolourous occurred at least 16 times with five reversals (**Figure 3.3**).

Pattern has a low phylogenetic signal (λ=0.25, $p<0.01$) indicating a near Brownian model of evolution. But life history trait has a stronger phylogenetic signal (λ=1, $p<0.01$) than colour (λ=0.80, $p<0.01$). On nodes reconstructed to be annual, four transitions happened from AR to white/yellow whereas for the perennial nodes, 10 such transitions happened (**Figure 3.2**). Without taking phylogenetic dependence into consideration, Pagel's correlation analysis rejected the independent evolution of colour and life cycle trait (4-parameter logL = -80.81, 8-parameter logL = -78.32, difference = 2.49, $p<0.05$, 500 simulations). However, the correlation coefficient under phylogenetic independent

contrast (PIC) is only 0.037 ($p<0.05$). And therefore, no correlation was detected between the life cycle trait and colour trait. Since life cycle is closed related to breeding mechanism (Brahim et al., 2001), the finding infers that allogamous perennials were not found to be more anthocyanin-rich than the autogamous annuals.

3.4 Discussion

3.4.1 Justification for ancestral reconstruction based on BEAST tree

Species placement discrepancies between this and previous studies are largely due to gene choice, reconstruction methods and taxon sampling.

For gene choice, I have the advantage of using existing data that is executable for coalescent model analysis in *BEAST, while previous studies used maximum parsimony (Asmussen and Liston, 1998, Kenicer et al., 2005) and maximum likelihood (Schaefer et al., 2012) methods on concatenated datasets The multi-species coalescent model has the advantage of taking into account the different genealogical history of unlinked loci that frequently lead to cyto-nuclear discordance (Soltis and Kuzoff, 1995, Renoult et al., 2009). For this dataset of 105 taxa, the species phylogeny was inferred using an extensive analysis of 800 million MCMC iterations. Unfortunately, the same approach would have been computationally challenging for Schaefer et al.'s much larger dataset of over 300 species (not only *Lathyrus*).

However, the multi-species coalescent approach assumes incomplete lineage sorting as the only factor for gene tree discordance. In reality, gene tree discordance can be due to other factors such as alignment problems or hybridization events (Degnan and Rosenberg, 2009). Hybridization may be particularly problematic in recently diverged groups. The origin of the *Lathyrus* crown group has been estimated at c. 5.4-6.3 Ma using absolute ITS substitution rates based on published data of other legumes with similar life history (Kenicer et al., 2005), or 9 - 5 Ma using an uncorrelated relaxed clock model, based on four secondary calibration points from Lavin et al. (2005), with a Bayesian time estimate (Schaefer et al., 2012). Hybrids among extant *Lathyrus* species are rare, and often sterile, although sometimes hybrids between very closely related species are successful (Kearney, 1993). Therefore, the assumption of incomplete lineage sorting, rather than hybridization, as the dominant phylogenetic problem is reasonable.

Lastly, the multi-species coalescent method is most accurate with numerous unlinked loci, dense taxon sampling and multiple samples per species (Zwickl and Hillis, 2002, Nabhan and Sarkar, 2011, Heled and Drummond, 2010). Recent *Lathyrus* studies have improved sampling density from earlier studies which had at most 42 (Asmussen and Liston, 1998) or 53 species (Kenicer et al. 2005). I included 92 out of the total of 160 *Lathyrus* taxa, representing 7 out of the 11 sections, and spanning the regions of Laurasia, Africa (*L. hygrophilus*) and the Americas, a wide sampling with moderate density. However, a 60% inclusion is not a full representation of the entire genus. Many recent splits all received low support (Figure 1). The coalescent model requires taxa to be represented in both linkage groups, therefore, some species with sequences from only one linkage group has to be excluded from my analysis. However, using ML reconstruction method, Schaefer et al. (2012) were able to include 110 *Lathyrus* species, an enhanced scale of sampling although my clades with low support failed to receive higher support value in their analysis.

The ancestral state seemed to be independent of tree search method. As a test, I reconstructed the phylogeny for the 105 taxa using a concatenated chloroplast and nuclear dataset with ML method. The resulting ancestral colour and pattern were estimated to be pigmented and concolourous with over 75% probability (data not shown). This indicates that reconstruction of character evolution for colour and pattern is relatively robust to minor differences in tree topology caused by differences in tree search algorithms. However, future study could aim for a wider sampling of the genus with more nuclear genes using coalescent approach.

3.4.2 Lack of association between anthocyanin-pigmented form and life history

Floral colour is evolutionarily labile in diverse genera and families (Rausher, 2008, McEwen and Vamosi, 2010). Here however, colour (versus the lack of) showed a high phylogenetic signal which could be due to the fact that over 75% of the *Lathyrus-Pisum* clade is composed of coloured taxa. And over 50% of the unpigmented taxa were grouped together in series *Lutei*. Overall, colour (pink, blue, purple, yellow) could still be a labile trait in *Lathyrus,* if coded as multiple traits instead of a binary trait. It is the presence of absence of anthocyanin that has a strong phylogenetic basis. As for life history trait, the biggest section *Orobus* is composed of perennial members only. And this life history strategy had evolved once in *Orobus* from the annual ancestral state.

Self-pollination is generally viewed as an adaptive strategy to the inconsistency or complete absence of pollinators, to provide reproductive assurance (Kalisz et al., 2004). Annual *Lathyrus* are facultatively autogamous (Brahim et al., 2001, Brahim et al., 2002), whereby a plant is capable of self-pollination in the absence of pollinators but practices outcrossing when pollinators are available. Facultative autogamy is a more flexible breeding system than obligate autogamy – when a plant is unable to receive foreign pollen hence can only reproduce by self-pollination. Geitonogamy (where a flower is fertilized by pollen from a different flower of the same plant) is sometimes treated as a form of autogamy as it is genetically identical to "self-pollinating autogamy" but it is functionally outcrossing since it is mediated by pollinators. On the other hand, perennial *Lathyrus* are generally incapable of self-pollination (Brahim et al., 2001), and are either outcrossing (allogamous/xenogamous) or geitonogamous, in both cases would require pollination by insects. For instance *L. latifolius* relies on its natural pollinator leafcutter bees (Westerkamp, 1993) and sets no seed in a bagging experiment (Brahim et al., 2001).

It was expected that functional autogamy should lower the selection pressure on floral display in annual *Lathyrus*. In support of this a reduced number of flowers per inflorescence was found in annual *Lathyrus (Kupicha, 1983)*. If colour is adaptive in *Lathyrus*, the annuals might be duller in colour due to relaxed selection. However, the lack of correlation when phylogenetic independence was taken into consideration showed no difference in floral colour display between the autogamous and allogamous species. Bees can detect ultraviolet light, which are often present in yellow and white species (Chittka et al., 1994). Therefore, a purely human-perspective of the colour is not necessary a true indicator of floral attractiveness (details in **3.3.4**).

3.4.3 Polarity of trait evolution

Loss of floral colour in *Lathyrus* is asymmetrical (or unidirectional) with no reversal inferred. This agrees with the general trend that changes from blue species to red, yellow or white species happen more often than the reverse, as observed in many genera such as *Ipomoea* (Rausher, 2008), *Penstemon* (Wilson et al., 2004), *Iochroma* (Smith et al., 2013) and *Sinningieae* (Perret et al., 2003). This asymmetry can be explained by the fact that a loss of pigments is in general due to a loss-of-function mutation, which is irreversible and can be reinforced if two or more mutations are required for a transition (Rausher, 2006).

However, reversals happened for both life history and pattern. In *Arabidopsis*, the annual or perennial trait is controlled by two genes, having redundant roles in flowering time (Melzer et al., 2008). Similarly, bicolour pattern is controlled together with a dark seed coat (via a pleiotropic effect) as a result of differential anthocyanin expression in the runner bean *Phaseolus*, by a transcription factor *bic* (Bassett and Miklas, 2007). If the genetic control of those two traits in *Lathyrus* is similar to those in *Arabidopsis* and *Phaseolus*, the reversibility of the traits can only be explained by strong selection on the gain-of-function mutations. And this implies that bicolourous and perennial traits are adaptive in *Lathyrus*.

How about the shift in colour? Is it a result of adaptive or non-adaptive processes? It seems that even though colour shifts are in some systems regularly coupled with pollinator shifts, e.g. in *Aquilegia* (Hodges and Derieg, 2009) , there are other potential adaptive mechanisms for colour change that do not involve pollinators. Instead selection affecting floral colour can be indirect through linkage to pleiotropic vegetative traits, which may be common (Armbruster, 2002, Arista et al., 2013). There are numerous published studies associating shifts in floral colour with shifts in pollinator type (Bradshaw and Schemske, 2003, Altshuler, 2003, Hoballah et al., 2007b), but only in few cases adaptive floral colour change (Fenster et al., 2004) has been rigorously demonstrated, e.g. in *Antirrhinum* (Schwinn et al., 2006), *Mimulus* (Streisfeld and Kohn, 2007, Streisfeld and Kohn, 2005, Bradshaw and Schemske, 2003). Pollinators of *Lathyrus* include members of the order Hymenoptera such as *Apis mellifera* (honey bee), *Bombus* and *Megachile* (leafcutter bee) (Knuth et al., 1906) with a potentially bird-pollinated species *L. splendens.* It is not possible, so far at least, to speculate if pollinators exert selective pressure on the flower colour evolution in *Lathyrus*.

3.4.4 Spectral limitation of this study

In this study, colour is assigned using the human perception of colour (visible spectrum 420 – 800 nm), species that appear the same colour to human eyes might be different for insect pollinators whose spectrum range is different at 300–660 nm (Altshuler, 2003). Colour receptors (opsins) for many trichromatic hymenopterans peak in the UV, blue and green spectral regions (Chittka and Raine, 2006) whereas human opsins have peak absorptions at red, green and blue regions. Blue and UV appear as UV-blue or UV-violet to the bees whereas yellow or green appear as UV-purple. In addition, species such as *L. cicera* reflect both red and UV, hence it appears as red to human while UV to a bee (**Figure 3.4a**).

Loss of anthocyanin results in a change in the ratio of flavonol to anthocyanin. In contrast to anthocyanins which have absorption peaks at 510 – 530 nm, flavonols absorbs mostly in the UV region. This difference may result in a UV pattern if ratios are drastically different for adjacent floral parts. Hence, petals that appear concolourous to the human eye, may show a bicolourous UV pattern such as *L. graminifolius* (**Figure 3.4b**). I was unable to document the UV pattern for the other taxa but this would be an interesting subject for further work.

Table 3.1 GenBank accession numbers and traits for the taxa studied. Bi = bicolourous for banner and winged petals; con, concolourous; P, perennial, A, annual. Classification followed Asmussen and Liston (1998), distribution largely follows Schaefer et al. (2012).

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Orobus	*L. alpestris* (Waldst. &Kit.) Čelak	Europe	Purple	Con	P	AY839341	AY839409	AY839476	--	--	--
Orobus	*L. aureus* (Steven) Bornm.	Caucasus Mts.	Yellow	Con	P	AY839347	AY839415	AY839481	--	--	--
Orobus	*L. davidii* Hance	Asia	Yellow	Con	P	AY839350	AY839418	AY839492	GQ436352	HM026393	GU396758
Orobus	*L. delnorticus* Hitchc.	N. America	White	Con	P	AY839351	AY839419	--	--	--	
Orobus	*L. emodii* Fritsch	Asia	Yellow	Con	P	JX506068	--	--	--	JX505795	--
Orobus	*L. eucosmus* Butters & St. John	N. America	Pink	Bi	P	JX506069	JX505662	--			--
Orobus	*L. glandulosus* Broich	N. America	Purple	Bi	P	AY839355	AY839423	AY839493	--	--	--
Orobus	*L. gmelinii* Fritsch	China, Russia	Yellow	Con	P	AY839357	AY839475	AY839537	--	--	--
Orobus	*L. holochlorus* (Piper) Hitchc.	N. America	White	Con	P	AY839359	AY839426	AY839490	--	--	--
Orobus	*L. humilis* (Ser.) Spreng.	China, Russia	Purple	Bi	P	AY839360	AY839427	AY839494	--	--	--
Orobus	*L. incurvus* Willd.	E. Mediterranean, Asia Minor	Purple	Bi	P	JX506087	JX505668	JX505544	--	--	JX505930
Orobus	*L. japonicus* Willd.	Laurasia	Purple	Bi	P	AY839361	AY839428	AY839495	JN661183	JX505800	AB611010
Orobus	*L. jepsonii* E. Greene	N. America	Pink	Bi	P	AY839362	AY839429	AY839496	--	--	--
Orobus	*L. komarovii* Ohwi	China, Mongolia, Korea, Russia	Purple	Con	P	AY839363	AY839430	AY839478	--	--	--

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Orobus	*L. laevigatus* Gren.	E. Europe	Yellow	Con	P	AY839365	AY839431	AY839497	JX505481	--	JX505932
Orobus	*L. lanszwertii* Kellogg	N. America	White	Con	P	AY839366	AY839433	AY839499	--	EU025896	--
Orobus	*L. laetiflorus* Greene	N. America	Pink	Con	P	JX506092	JX505670	--	--	--	--
Orobus	*L. linifolius* (Reichard) Bässler	Europe	Pink	Con	P	AY839368	AY839435	AY839477	--	--	--
Orobus	*L. littoralis* (Nutt.) Endl.	West N. America	Purple	Bi	P	AY839369	AY839436	AY839500	--	--	--
Orobus	*L. luteus* Baker	Asia	Yellow	Con	P	JX506098			--	JX505802	JX505935
Orobus	*L. multiceps* Clos	Chile	Purple	Bi	P	AY839370	AY839437	AY839484	--	--	--
Orobus	*L. nevadensis* S. Watson	N. America	Purple	Con	P	AY839374	AY839441	AY839501	--	--	--
Orobus	*L. ochroleucus* Hook.	Canada	White	Con	P	JX506110	--	--	--	JX505806	JX505942
Orobus	*L. palustris* L.	Laurasia	Purple	Con	P	AY839379	AY839446	AY839502	JN890637	HM026396	--
Orobus	*L. palustris* ssp.*pilosus* Cham.	Asia	Purple	Con	P	JX506118	JX505679	JX505557	--	HM026396	--
Orobus	*L. palustris* var. *pilosus* Ledeb	Laurasia	Purple	Con	P	AY839380	AY839448	AY839504	--	--	--
Orobus	*L. pauciflorus* Fern.	N. America	Purple	Bi	P	JX506117	JX505678	JX505556	--	--	--
Orobus	*L. parvifolius* Watson	Mexico	Purple	Con	--	JX506116	--	--	--	--	JX505946
Orobus	*L. pisiformis* L.	Europe, N. Asia	Purple	Con	P	AY839382	AY839450	AY839491	--	--	--
Orobus	*L. polymorphus* Nutt.	N. America	Purple	Bi	P	JX506120	JX505681				--
Orobus	*L. polyphyllus* Torr. & A.Gray	N. America	Pink	Bi	P	AY839383	AY839449	AY839505	--	--	--
Orobus	*L. quinquenervius* (Miq.) Litv.	China, Japan, Korea, Russia	Purple	Con	P	AY839386	AY839453	--	--	JX505814	JX505950

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Orobus	*L. rigidus* T. White	N. America	White	Con	P	AY839387	AY839454	AY839506	--	--	--
Orobus	*L. splendens* Kellogg	N. America	Red	Con	P	AY839395	AY839462	AY839507	--	--	--
Orobus	*L. subandinus* Phil.	S. America	Purple	Bi	P	AY839396	AY839463	AY839483	--	--	JX505958
Orobus	*L. transsylvanicus* Rchb.f.	E. Central Europe	Yellow	Con	P	AY839400	AY839467	AY839509	--	--	--
Orobus	*L. vaniotii* Leveille	Asia	Blue	Con	P	AY839402	AY839469	AY839511	--	HM026398	--
Orobus	*L. venetus* (Mill.) Wohlf.	Russia	Pink	Con	P	JX506153	--	JX505579	--	--	--
Orobus	*L. venosus* Muhl. ex Willd.	N. America	Purple	Bi	P	JX506154	JX505704	JX505580	JX505484	JX505823	JX505961
Orobus	*L. vernus* (L.) Bernh.	Eurasia	Purple	Bi	P	AY839404	AY839470	AY839480	--	--	--
Orobus	*L. vestitus* Nutt.	N. America	Pink	Bi	P	AY839405	AY839472	AY839510	--	AF522088	--
Orobus	*L. zionis* C.L. Hitchc.	N. America	Pink	Con	P	JX506156	JX505707	JX505582	--	JX505824	--
Lathyrus	*L. amphicarpos* L.	Mediterranean	Red	Con	A	JX506047	JX505647	--	--	--	--
Lathyrus	*L. annuus* L.	Mediterranean	Yellow	Con	A	AY839344	AY839412	AY839525	--	--	--
Lathyrus	*L. belinensis* Maxted & Goyder	Turkey	Red	Bi	A	JX506062	--	--	JX505476	JX505792	JX505921
Lathyrus	*L. blepharicarpus* Boiss.	E. Mediterranean	Orange	Con	A	JX506063	JX505660	JX505536	--	--	--
Lathyrus	*L. chloranthus* Boiss.	Syria	Yellow	Con	A				--	--	--
Lathyrus	*L. chrysanthus* Boiss.	Turkey	Yellow	Con	A				--	--	--
Lathyrus	*L. cicera* L.	Mediterranean	Red	Con	A	AY839348	AY839416	AY839518	--	--	--
Lathyrus	*L. crassipes* Phil.	S. America	Purple	Con	P	JX506066	--	--	--	--	JX505922
Lathyrus	*L. grandiflorus* Sibth. & Sm.	Mediterranean	Pink	Bi	A	JX506075	--	JX505540	--	JX505797	JX505924

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Lathyrus	*L. heterophyllus* L.	Europe	Pink	Con	P	AY839358	AY839425	AY839532	--	--	--
Lathyrus	*L. hierosolymitanus* Boiss.	Mediterranean	Yellow	Con	A	JX506079	--	JX505542	JX505477	--	JX505925
Lathyrus	*L. hirsutus* L.	Mediterranean, Europe, Macaronesia	Purple	Bi	A	JX506080	--	--	JN661182	JX505798	JX505926
Lathyrus	*L. latifolius* L.	Europe	Pink	Con	P	AF335219	--	--	HM029364	AF522085	--
Lathyrus	*L. mulkak* Lipsky	Asia	Pink	Con	A	JX506105	--	JX505548	--	--	JX505940
Lathyrus	*L. odoratus* L.	Sicily	Purple	Bi	A	AY839377	AY839474	AY839533	--	--	--
Oronbon	*L. roseus* Steven	Caucasus Mts.	Red	Con	P	JX506131	JX505687	JX505563	--	JX505815	JX505951
Lathyrus	*L. rotundifolius* Willd.	Mediterranean, Asia	Red	Con	A	AY839388	AY839455	AY839535	--	JX505816	JX505952
Lathyrus	*L. sativus* L.	Eurasia, N. Africa	Blue	Con	A	AY839389	AY839456	AY839517	--	AF522086	--
Lathyrus	*L. sylvestris* L.	Europe	Pink	Con	P	AY839398	AY839465	AY839523	--	--	--
Lathyrus	*L. tuberosus* L.	Mediterranean, Europe	Pink	Con	P	JX506149	JX505701	JX505576	JN661188	--	--
Lathyrus	*L. tukhtensis* Czeczott	Turkey	Blue	Con	P	JX506151	--	JX505578	--	--	JX505960
Lathyrostylis	*L. bauhinii* Genty	Alps, Pyrennes	Purple	Bi	P	JX506061	JX505659	JX505535	JX505475	--	--
Lathyrostylis	*L. digitatus* (Bieb.) Fiori & Poal.	Mediterranean	Purple	Bi	P	AY839353	AY839421	AY839514	--	--	--
Lathyrostylis	*L. filiformis* (Lam.) Gay	S. Europe	Purple	Bi	P	AY839354	AY839422	AY839536	--	--	--
Lathyrostylis	*L. pallescens* (Bieb.) Koch	Caucasus Mts.	White	Con	P	AY839378	AY839445	AY839515	--	--	--
Lathyrostylis	*L. pannonicus* (Jacq.) Garcke	Europe	White	Con	P	JX506113	JX505677	JX505554	--	JX505808	--

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Lathyrostylis	*L. spathulatus* Čelak	Asia Minor	Purple	Bi	P	AY839392	AY839459	AY839513	--	--	--
Lathyrostylis	*L. tingitanus* L.	Mediterranean, Macaronesia	Red	Con	A	AY839399	AY839466	AY839519	JN661187	JX505822	JX505959
Pratensis	*L. laxiflorus* Kuntze	Mediterranean, Asia	Purple	Bi	P	AY839367	AY839434	AY839482	--	--	JX505933
Pratensis	*L. pratensis* L.	Eurasia	Yellow	Con	P	JX506122	AY839451	JX505559	JN661185	JX505810	JX505947
Aphaca	*L. aphaca* L.	Mediterranean, Macaronesia	Yellow	Con	A	AY839345	AY839413	AY839489	JX505474	JX505788	JX505918
Clymenum	*L. gloeospermus* Warb. & Eig	E. Mediterranean	White	Con	A	AY839356	AY839424	AY839527	--	JX505796	--
Clymenum	*L. articulatus* L.	W. Mediterranean	Red	Bi	A	JX506058	JX505656	JX505532	JN661181	JX505791	JX505920
Clymenum	*L. clymenum* L.	Mediterranean	Red	Bi	A	AY839346	AY839417	AY839529	--	JX505793	--
Clymenum	*L. ochrus* (L.) DC	Canary Islands	White	Con	A	AY839376	AY839444	--	JN661184	JX505807	JX505943
Orobastrum	*L. setifolius* L.	S. Europe	Red	Con	A	AY839391	AY839458	AY839516	--	JX505818	JX505954
Linearicarpus	*L. angulatus*_Kenicer L.	Mediterranean, Macaronesia	Purple	Bi	A	AY839343	AY839411	AY839522	--	--	--
Linearicarpus	*L. angulatus*_Schaefer L.					JX506049	JX505649	JX505523	JN661180	JX505787	JX505917
Linearicarpus	*L. hygrophilus* Taubert	Tropical Africa	Purple	--	A	JX506082	--	--	JX505478	--	JX505927
Linearicarpus	*L. inconspicuus* L.	Asia	Purple	Bi	A	JX506083	--	JX505543	JX505479	--	JX505928
Linearicarpus	*L. sphaericus* Retz.	Mediterranean, Asia	Red	Con	A	AY839393	AY839461	AY839512	JN661186	JX505820	JX505956
Nissolia	*L. nissolia* L.	S. Europe, Asia	Red	Con	A	AY839375	AY839443	AY839520	--	JX505805	--
Neurolobus	*L. neurolobus* Boiss. & Heldr.	Crete (Greece)	Purple	Bi	P	AY839373	AY839440	AY839521	--	JX505804	--
Notolathyrus	*L. nervosus* Lam.	S. America	Purple	Bi	P	AY839372	AY839439	AY839486	JX505482	JX505803	JX505941

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Ingroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
Notolathyrus	*L. magellanicus* Lam.	S. America	Purple	Bi	P	AY839371	AY839438	AY839485	--	--	JX505937
Notolathyrus	*L. paranensis* Burkart	S. America	Blue	Bi	P	JX506114	--	--	--	--	JX505944
Notolathyrus	*L. parodii* Burkart	S. America	Purple	Con	P	JX506115	--	--	--	--	JX505945
Notolathyrus	*L. pubescens* Hook. & Arn.	S. America	Purple	Bi	P	AY839385	AY839452	--	JX505483	JX505812	JX505949
--	*L. meridensis* Pittier	Columbia	Pink	--	P	JX506103	--	--	--	--	JX505939
--	*L. linnaei* Rouy	Asia	--	--	P	JX506096	--	--	--	JX505801	JX505934
	Pisum fulvum Sibth. & Sm.	Asia Minor	Orange	Con	A	AB546787	AB546803	AB546819	--	--	FR856873
	Pisum sativum L.	Mediterranean, Asia Minor	Purple	Bi	A	AY839340	AY839473	AY839526	--	JX505828	--
Outgroups											
	Cicer canariense Santos & Lewis	Canary Islands	Purple	Con	P	AJ639941	AB117650	AB257390	--	AF522079	--
	Cicer kermanense Bornm.	Asia	Pink	Con	P	AB198915	AB117657	AB257386	--	AB198884	--
	Cicer microphyllum Benth.	Asia	Purple	Con	P	AJ639943	AB117659	AB257382	--	AB198886	--
	Cicer multijugum Maesen	Asia Minor	Purple	Con	P	AJ639944	AB117660	AB257392	--	AB198888	--
	Lens culinaris Medik.	Asia Minor	Blue	Con	A	AB546788	AB546804	AB546820	JN661189	JX505825	JX505962
	Lens ervoides (Brign.) Grande	S. Europe	Blue	Bi	A	AB546791	AB546823	AB546807	--	AF522089	--
	Lens nigricans Webb. & Berthel.	Mediterranean, Asia Minor	Purple	Con	A	AJ441070	--	JX505584	JX505485	--	JX505963

Section	**Taxa**	**Distribution**	**Colour**	**Pattern**	**Life forms**	**Accession**					
Outgroups						***ITS***	***trnL-F***	***trnS-G***	***rbcL***	***matK***	***psbA-trnH***
	Lens orientalis Hand.-Mazz.	Russia	Blue	Con	A	AB546789	AB546805	AB546821	--	--	--
	Vicia cracca L.	Europe	Purple	Con	P	AY839339	AY839406	AY839530	JN661198	--	JX505993
	Vicia nipponica Matsum.	Japan	Purple	Con	P	AY839338	AY839407	AY839534	--	HM026404	--
	Vicia unijuga Braun	Asia	Purple	Con	P	AY839337	AY839408	AY839531	--	HM026402	--

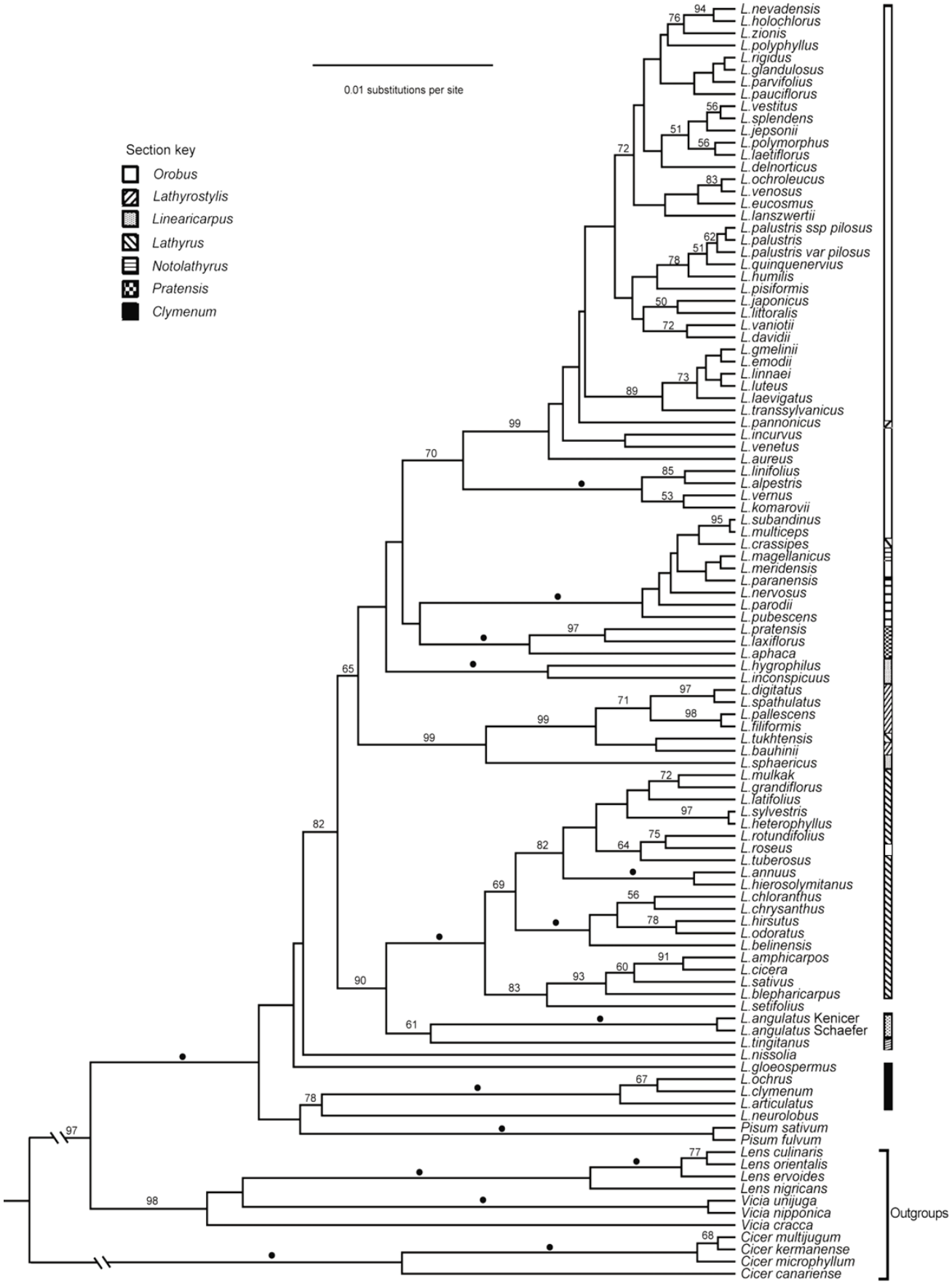
0.01 substitutions per site
Section key
Orobus
Lathyrostylis
Linearicarpus
Lathyrus
Notolathyrus
Pratensis
Clymenum
L.nevadensis
L.holochlorus
L.zionis
L.polyphyllus
L.rigidus
L.glandulosus
L.parvifolius
L.pauciflorus
L.vestitus
L.splendens
L.jepsonii
L.polymorphus
L.laetiflorus
L.delnorticus
L.ochroleucus
L.venosus
L.eucosmus
L.lanszwertii
L.palustris ssp pilosus
L.palustris
L.palustris var pilosus
L.quinquenervius
L.humilis
L.pisiformis
L.japonicus
L.littoralis
L.vaniotii
L.davidii
L.gmelinii
L.emodii
L.linnaei
L.luteus
L.laevigatus
L.transsylvanicus
L.pannonicus
L.incurvus
L.venetus
L.aureus
L.linifolius
L.alpestris
L.vernus
L.komarovii
L.subandinus
L.multiceps
L.crassipes
L.magellanicus
L.meridensis
L.paranensis
L.nervosus
L.parodii
L.pubescens
L.pratensis
L.laxiflorus
L.aphaca
L.hygrophilus
L.inconspicuus
L.digitatus
L.spathulatus
L.pallescens
L.filiformis
L.tukhtensis
L.bauhinii
L.sphaericus
L.mulkak
L.grandiflorus
L.latifolius
L.sylvestris
L.heterophyllus
L.rotundifolius
L.roseus
L.tuberosus
L.annuus
L.hierosolymitanus
L.chloranthus
L.chrysanthus
L.hirsutus
L.odoratus
L.belinensis
L.amphicarpos
L.cicera
L.sativus
L.blepharicarpus
L.setifolius
L.angulatus Kenicer
L.angulatus Schaefer
L.tingitanus
L.nissolia
L.gloeospermus
L.ochrus
L.clymenum
L.articulatus
L.neurolobus
Pisum sativum
Pisum fulvum
Lens culinaris
Lens orientalis
Lens ervoides
Lens nigricans
Vicia unijuga
Vicia nipponica
Vicia cracca
Cicer multijugum
Cicer kermanense
Cicer microphyllum
Cicer canariense
Outgroups

Figure 3.1 Bayesian multi-species coalescent estimate of species phylogeny based on combined plastid sequences and nuclear ITS datasets, as two independently evolving linkage groups. Posterior probabilities (PP) greater than 50% were shown. ● represents 100% pp.

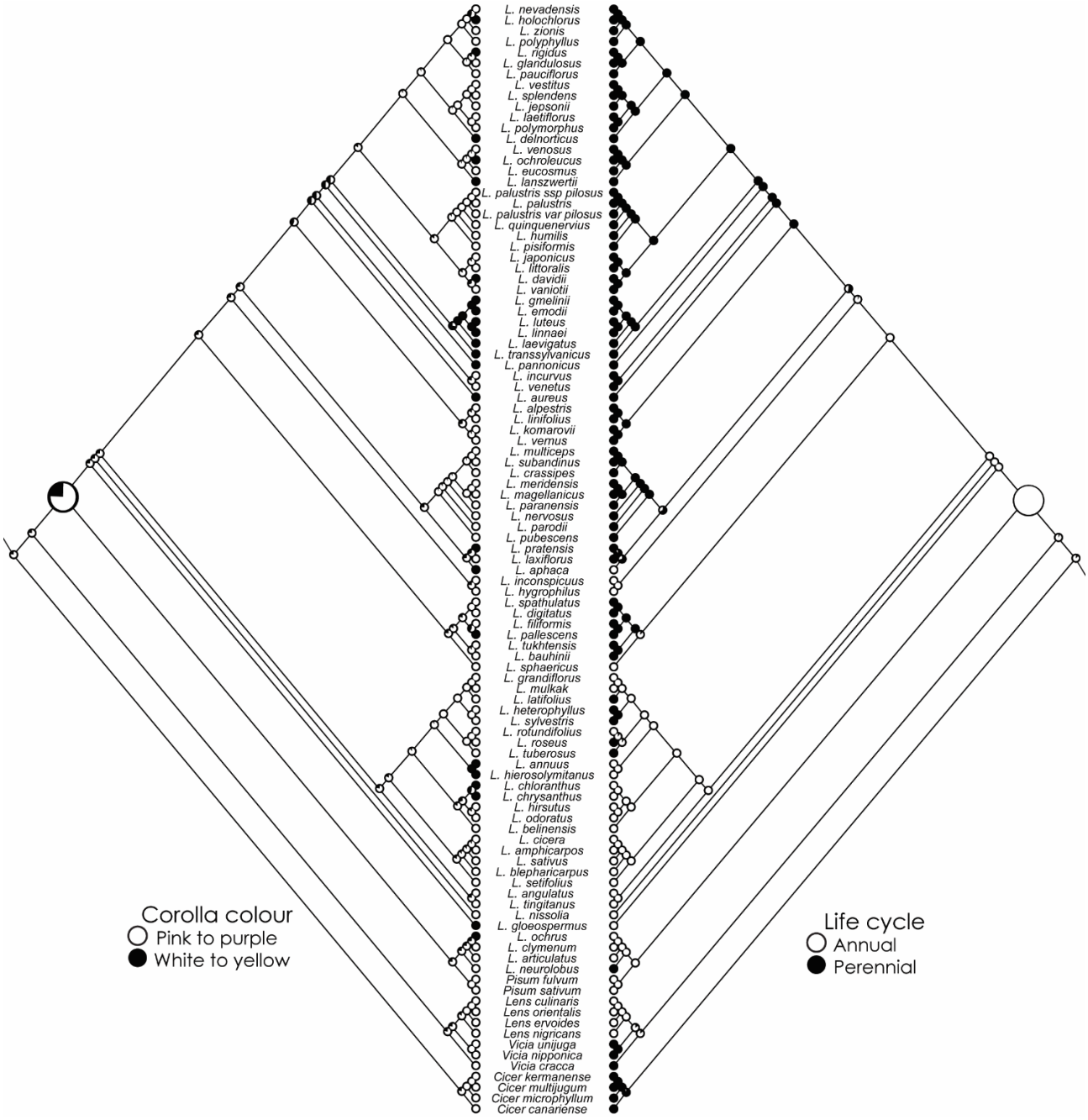

Figure 3.2 Ancestral colour and life cycle reconstruction on the maximum likelihood tree, under the Markov *k*-state two-parameter model. Circles at tips indicate character states and pie charts at nodes indicate ancestral character states. Left, floral colour: white = white to yellow, black = pink to purple. Right, life cycle: white = perennial, black = annual.

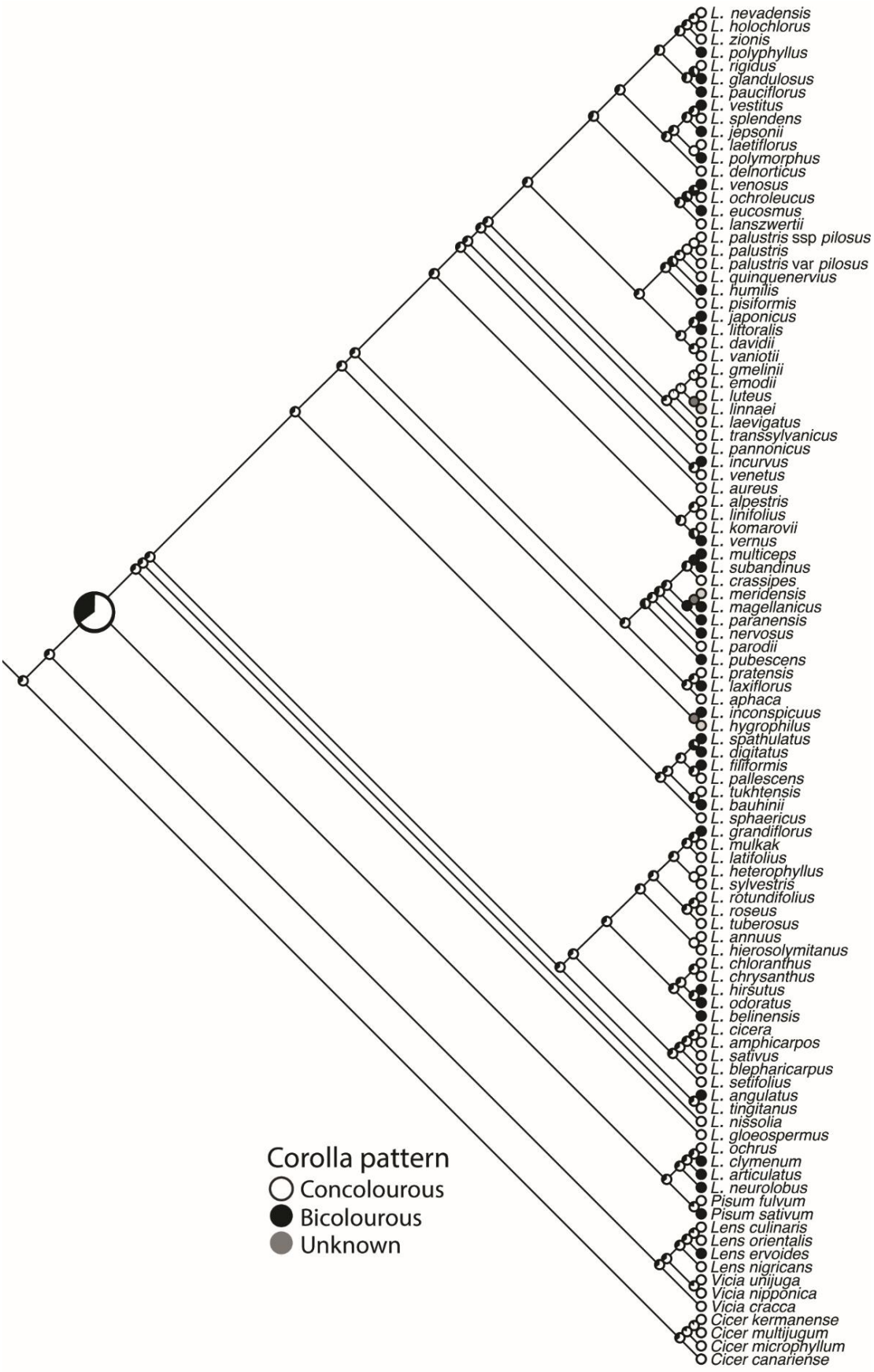
L. nevadensis
L. holochlorus
L. zionis
L. polyphyllus
L. rigidus
L. glandulosus
L. pauciflorus
L. vestitus
L. splendens
L. jepsonii
L. laetiflorus
L. polymorphus
L. delnorticus
L. venosus
L. ochroleucus
L. eucosmus
L. lanszwertii
L. palustris ssp pilosus
L. palustris
L. palustris var pilosus
L. quinquenervius
L. humilis
L. pisiformis
L. japonicus
L. littoralis
L. davidii
L. vaniotii
L. gmelinii
L. emodii
L. luteus
L. linnaei
L. laevigatus
L. transsylvanicus
L. pannonicus
L. incurvus
L. venetus
L. aureus
L. alpestris
L. linifolius
L. komarovii
L. vernus
L. multiceps
L. subandinus
L. crassipes
L. meridensis
L. magellanicus
L. paranensis
L. nervosus
L. parodii
L. pubescens
L. pratensis
L. laxiflorus
L. aphaca
L. inconspicuus
L. hygrophilus
L. spathulatus
L. digitatus
L. filiformis
L. pallescens
L. tukhtensis
L. bauhinii
L. sphaericus
L. grandiflorus
L. mulkak
L. latifolius
L. heterophyllus
L. sylvestris
L. rotundifolius
L. roseus
L. tuberosus
L. annuus
L. hierosolymitanus
L. chloranthus
L. chrysanthus
L. hirsutus
L. odoratus
L. belinensis
L. cicera
L. amphicarpos
L. sativus
L. blepharicarpus
L. setifolius
L. angulatus
L. tingitanus
L. nissolia
L. gloeospermus
L. ochrus
L. clymenum
L. articulatus
L. neurolobus
Pisum fulvum
Pisum sativum
Lens culinaris
Lens orientalis
Lens ervoides
Lens nigricans
Vicia unijuga
Vicia nipponica
Vicia cracca
Cicer kermanense
Cicer multijugum
Cicer microphyllum
Cicer canariense
Corolla pattern
Concolourous
Bicolourous
Unknown

Figure 3.3 Ancestral reconstruction for colour pattern on the maximum likelihood tree, under the Markov *k*-state two-parameter model. Circles at tips indicate character states and pie charts at nodes indicate ancestral character states. White: concolourous; black: bicolourous; grey: missing state.

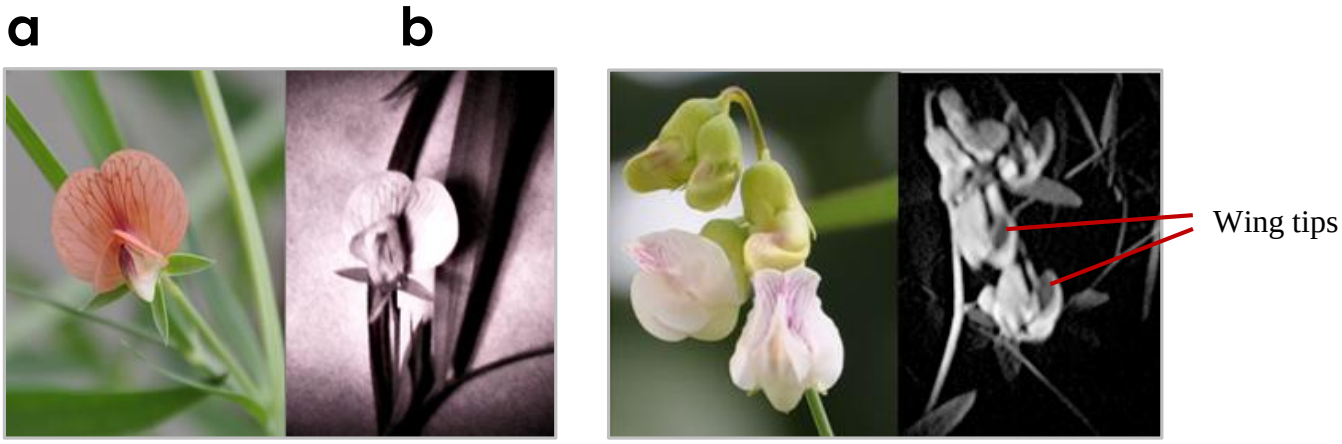

Figure 3.4 UV-reflecting patterns in two concolourous *Lathyrus* species. (a) *L. cicera* reflects UV in all parts, whereas **(b)** floral parts of *L. graminifolius* reflect UV differentially with the tip of the wings strongly absorbing at the UV region.

Chapter 4. The Genetic Basis for Parallel Evolution of 'White Shifts' in *Lathyrus*

4.1 Introduction

The causes and consequences of character transitions is an important issue in evolutionary biology. Similar character state transitions may indicate functional convergence (Barrett, 2008). I use 'white shifts' to imply transitions to white, cream and yellow-flowered taxa from their coloured relatives. Loss of colour may have fitness consequences due to either biotic (e.g pollination, herbivory) (Strauss and Whittall, 2006) or abiotic (stress from heat, drought and high UV exposure) mechanisms. Multiple 'white shifts' provide examples to investigate whether similar transitions have similar molecular and developmental mechanisms. Note that the 'white shifts' is an evolutionary transition instead of an ontogenetic change (such as those detailed by (Weiss, 1995).

4.1.1 Species used in this study

In the genus *Lathyrus* as a whole, loss of anthocyanin happened at least 16 times with no reversals (**Chapter 3**). This is therefore an excellent system in which to study the mechanism of replicated evolution although it is not yet known whether any of the colour switches have been adaptive. This 'white shift' leads to an albino phenotype, at least in terms of anthocyanins because other pigments maybe present such as flavonols in the sap, or carotenoids in the plastids. Therefore yellow species are included here in the 'white shifts' category, their yellow colour can be either due to plastid or soluble non-anthocyanin pigments (Onslow, 2014).

The genus *Lathyrus* have zygomorphic (bilaterally symmetrical) flowers, which gives a complex outer contour to enable efficient pollinator recognition. Many species also possess bilaterally symmetrical nectar guides forming an internal patterning. Presumably for efficient pollination, the colour of the nectar guide often contrasts with that of the petal. The evolutionary loss of pigment in this thesis refers to the main part of the flower, and in some cases anthocyanin-based nectar guides may still be present. I investigated the genetic basis for 'white shifts' that have originated both naturally and under cultivation. This has two major advantages: **1)** selective breeding on certain genetic traits makes the

genetics easier to study and **2)** artificial and natural selection may impose different selective pressure on species and therefore provides an excellent comparison to contrast gene usage for particular phenotypes under different selection modes.

The grass pea *L. sativus* is a widely cultivated crop, and it is among the first crops to have been domesticated, as early as 6000 BC (Kislev, 1989). It adapts well to arid regions of the world which are often poor and vulnerable to climate change and its associated natural disasters. Growing *L. sativus* in those regions can therefore help to alleviate malnutrition. The wild type cultivar *L. sativus* var. *azureus* (*L. sativus* in this thesis), also colloquially called 'King Tut blue sweet pea', is rumored to have been found in the tomb of the King Tutankhamum ca. 1300 BC. The albino cultivar, *L. sativus* var. *albus*, also has a long cultivation history. Another cultivated species studied here is the sweet pea *L. odoratus*, a popular ornamental around the world with its first recorded cultivation dating back to the 17th century. Genetics of floral colour in the sweet pea has been extensively studied from the beginning of the 20th century (Bateson et al., 1906, Punnett, 1923). Two loci can independently cause 'white shifts' in *L. odoratus* (Beale et al., 1939) and I have investigated one of these loci here.

4.1.2 Genetic control of floral colour

Structural genes in the anthocyanin pathway are grouped into early (*CHS, CHI, F3H*) and late genes (*DFR, ANS, OMT* and *UFGT*). It's commonly considered that the early and late genes are separately co-regulated (Wei et al., 2011). It is true that in some cases, e.g *Petunia*, most of the early genes (*CHS-A, CHI* and *F3H*) are regulated independently from the late genes (*DFR* onwards) plus *F3'5'H* and *CHS-J* (Gerats and Strommer, 2009). However this is not always the case and co-regulation of the early and late genes has been reported in *Zea mays* (Uimari and Strommer, 1997, Mol et al., 1998, Quattrocchio et al., 1993, Quattrocchio et al., 1998), *Ipomoea* (Durbin et al., 2003, Clegg and Durbin, 2003), *Mimulus* (Streisfeld and Rausher, 2009) and *Antirrhinum* (Schwinn et al., 2006).

The primary regulators of the anthocyanin pathway are the 'MBW complex', comprising transcription factors (TFs) of three families: R2B2-**M**YB, basic Helix-Loop-Helix (**b**HLH) and **W**D40 repeat proteins. These interact with each other and activate the structural genes by binding to their *cis*-

elements (Holton and Cornish, 1995, Ramsay and Glover, 2005, Albert et al., 2014). **Table 4.1** shows regulators of the anthocyanin pathway in some organisms. There is a high level of functional conservation through MBW genes in the angiosperms (Quattrocchio et al., 1993). For instance, homologues of *AN2* (a MYB gene) control flower colour shifts and pollinator preference in *Mimulus* (Hoballah et al., 2007a) and a single base substitution in *AN1* (an unrelated bHLH gene) led to the white *Pisum* mutant that Mendel studied 150 years ago (Hellens et al., 2010).

Changes in floral pigments are often due to few genes of large effects (Quattrocchio et al., 1999, Bradshaw and Schemske, 2003, Schwinn et al., 2006, Des Marais and Rausher, 2010). The 'white shift' is controlled by a single locus in *Petunia* (Quattrocchio et al., 1999) and two loci in *Antirrhinum (Schwinn et al., 2006)*, both TFs of the anthocyanin pathway. Examining whether, and how, this limited selection of genes is reused in replicated evolution can help us assess how predictable evolution is.

4.1.3 Evolution of floral colour change in *Lathyrus*

According to early studies (Bateson et al., 1906, Punnett, 1923), the white colour in *L. odoratus* was found to have originated twice with two independent mutations, which early geneticists distinguished as two 'factors' (C white or R white). However, one of these factors is apparently rare as most white sweet peas, when crossed with each other, segregate as a single locus (Fleming, 1925) and I did not find a genetically different white cultivar. The locus investigated in this thesis corresponds to R-white. C-white was therefore not investigated.

Morphological similarity can occur through several mechanisms, which may involve homology [synapomorphy (shared derived characters) or symplesiomorphy (maintenance of an ancestral state)] or homoplasy. Repeated colour transitions can be achieved through mechanism of **(1) convergent evolution**: Species with different ancestry acquiring analogous traits, e.g. the purple colour in yam (*Dioscorea alata*) and beetroot (*Beta vulgaris*) are due to anthocyanin and betalains respectively; **(2) parallel evolution** Related species evolving an independently developed but similar trait (Ostevik et al., 2012, Stern, 2013); and **(3) reversal** to the ancestral state (Wake, 1991, Cronk, 2009).

A loss of floral pigments may be examined against this evolutionary framework. In the genus *Lathyrus*, the ancestral state was anthocyanin-pigmented (**Chapter 3**), so white flowers are not a symplesiomorphy or reversal (although it is possible that at a deeper level, ancestral colour can still be white). The loss of pigments must therefore be due to a loss of anthocyanins from the flavonoid pathway (**Figure 4.1**) (Pecket, 1960) Hence **parallel evolution** is the most suitable category for the replicated inactivation of the anthocyanin pathway.

In addition to parallel evolution at the phenotypic level, on a genetic level parallelism would mean the independent occurrences of the same mutation resulting in the similar morphology. Parallelism at both levels have been reported in *Ipomoea* where *F3'H* was down-regulated in the red flowers in two independent lineages (Des Marais and Rausher, 2010). In addition, genetic parallelism can also arise through genetic constraints in which only a single gene is the possible route by which to make an evolutionary step happen (Wake, 1991).

4.1.4 Objectives

The question raised in this chapter is whether the parallel 'white shifts' in *Lathyrus* are also due to parallel genetic evolution, involving an inactivation of the same gene. Using a candidate gene approach, I studied the genetic basis for the 'white shifts' in the *L. odoratus* white cultivar 'Mrs Collier' ('MC'). Then I tested if this same gene was responsible for another domesticated 'white shift', that of the grass pea *L. sativus* var. *albus.* Expression of *DFR* gene was further tested in several unpigmented *Lathyrus* species of wild origin to find out if phenotypic parallelism, under different types of selection, has a parallel genetic basis. Finally, it should be mentioned that, although it is not within the scope of this study, future work on pollinator preference in *Lathyrus* will be important to determine whether the 'white shifts' in wild *Lathyrus* is adaptive.

4.2 Materials and methods

4.2.1 Plant materials

Seeds were obtained from the US National Plant Germplasm System or commercial growers (**Table 4.3**). Plants were grown at the UBC horticultural greenhouse facility with ambient light and temperature, between March and July 2012. For pigment extraction, flowers were harvested at anthesis into silica gel. Banner and wings were separately analyzed for the bicolourous *L. odoratus*. Among the unpigmented *Lathyrus*, because some possess nectar guides in the banner (**Table 4.3**), banner and wings were separately analyzed for two of them (*L. aphaca* and *L. chrysanthus*) as a test. *L. chloranthus*, without visible nectar guide was also separately analyzed for banner and wings, as a reference. For all other species, floral parts were combined for analysis. Pigment profile for the red bicolourous species *L. clymenum* was obtained as a comparison. Each biological replicate contained a pooled sample of at least three flowers from one individual plant. Pigment was analyzed for three biological replicates per taxa.

For RNA extraction, fresh petals of *L. odoratus* were harvested one or two days before anthesis by flash freezing in liquid nitrogen. Petals from all other species were collected between the mature bud stage and one day into anthesis, since the buds were small and may not provide enough material. Floral tissues from at least three individual plants were pooled as one sample.

4.2.2 Pigment identification using an LC-UV-MS system

Pigments were analyzed at the aglycone level. To extract water-soluble pigments, a consistent 20 mg ml^{-1} petal to solvent concentration was made in a Greiner tube for all samples by weighing 20-30 mg of silica-dried petals and soak it in 1-1.5 ml solvent of methanol:H_2O:acetic acid (85:15:0.5, v/v/v) at 4°C overnight with occasional vigorous shaking. This extraction method seemed to be effective for the complete extraction of all of the flavonoids in *Lathyrus*. The solution was then centrifuged to retain the supernatant. The extract was acid hydrolyzed with 2 M HCl for 30mins in a water bath (80°C) to remove the sugar residues. Precipitates were quickly spun down and the supernatant was kept in the dark at -20°C until further processing.

For LC-MS analysis, an Agilent 1100 LC/MSD system was equipped with a diode array detector (DAD) to detect UV absorption. The system contains a high pressure binary pump to allow for the mixing of two solvents. A ZORBAX SB-C18 reverse phase column (4.6 × 50 mm, 1.8 μm, Agilent) was thermostatted at 70°C for optimal elution of compounds with minimal usage of solvents. The crude extract was eluted at a rate of 1 ml min^{-1} with 87% solvent A [H_2O/FA (folic acid), 98:2, v/v] and 13% solvent B (acetyl nitrate/FA, 98:2, v/v) for 4 min, followed by an elution gradient from 23% B for 2 min to 90% B for 1 min, and then back to 13% B. Technical quadruplicates with injection volume of 5 μl were performed for all samples. The eluate absorbance at 357—373 nm (for flavonol aglycones) and 510—530 nm (for anthocyanidins) was determined using DAD. Accurate mass of the eluent was then determined by mass spectrometry (MS). Anthocyanidins and flavonol aglycones were examined at the positive and negative ionization modes respectively, following traditions. Peak identification was made by overlaying the MS chromatogram onto the LC chromatogram of a given sample, and pigment was assigned to each sample based on the retention time (T_R) and mass to charge ratio (m/z) of the reference standards and online database (massbank.jp). The criteria for pigment assignment were listed in **Table 4.4**. The relative abundance of an identified pigment was determined by calculating the total area under its peak absorption.

4.2.3 RNA extraction, transcriptome sequencing and RT-PCR

RNA was extracted using PureLink Plant RNA Reagent (Invitrogen) following manufacturer's protocol. Total RNA was treated with DNase twice using a Turbo DNA-free kit (Ambion). The first-strand cDNA was then synthesized from the total RNA (500 ng) with H Minus M-MLV reverse transcriptase (Fermentas) using supplied oligo$(dT)_{18}$ primer and following the manufacturer's protocol.

The *L. odoratus* transcriptome library was obtained as follows: flowers at mature bud stage of several individuals were flash frozen. RNA was extracted from the pooled samples and the quality was determined using a 2100 Bioanalyzer (Agilent). The RNA was then submitted for whole-transcriptome sequencing at Cofactor genomics (St Louis, Mo.) using the Illumina HiSeq platform. Saemi Sveinsson from the Cronk lab performed the *de novo* assembly using the Trinity assembly programme.

Gene expression tested includes structural gene *DFR* and *ANS*, transcription factors *AN1*, *ROI*, *AN2*, *AN4* and *AN11* homologues. To design *Lathyrus*-specific gene primers, I obtained the CDS of a few ABP genes by blasting the *L. odoratus* transcriptome library against a *Petunia* library. Contigs with the highest identity scores to *Petunia* homologues were then separately blasted in GenBank for verification. For all genes, primers spanned at least two exons to eliminate false signals given by DNA contamination. Each qualitative RT-PCR amplification used a reaction mixture containing 0.5 mM of each primer, 10 × PCR buffer (Fermentas), 2.0 mM $MgCl_2$, 0.25 mM of each dNTP, and 5 unit of recombinant Taq polymerase (Fermentas). PCR was performed for 30 cycles with each cycle condition: 30s at 94°C, 30s at the optimized annealing temperature, 1min at 72°C. PCR products were run on 1.5% agarose gels (0.5×TAE buffer) with 3 µl GelRed (Biotium) for band separation and visualization. If required for sequencing, PCR product was purified using a QIAquick Gel Extraction Kit following the manufacturer's instructions and sent for Sanger sequencing using an ABI sequencer at the NAPS facility at UBC.

Quantitative RT-PCR was performed with a Biorad iCycler iQ real-time PCR machine, following the manufacturer's protocol. Real-time RT-PCR results may be affected by primer specificity. To eliminate this possibility, I used primers that were specific to the individual species by first sequencing the *DFR* fragments using *L. odoratus* primer pairs. Each reaction included a 20 µl mixture of cDNA (50 ng), SssoFast EvaGreen supermix (10 µl), forward and reverse primer (5 pmol/µl each), and nuclease-free water. Each sample was performed with three technical replicates. All primers (**Table 4.2**) were tested first with qualitative PCR using templates of gDNA and RNA. Real-time PCR condition was: 30s at 95°C, followed by 40 cycles of 5s at 95°C and 20s at 51°C. Relative quantification used normalization against the reference gene *ACTIN2*.

4.2.4 Cosegregation analysis

A cosegregation analysis was done for *L. odoratus,* the wild type cultivar 'CP' was crossed with the white cultivar 'MC' to yield six F1 hybrids, which were allowed to self-pollinate. In the original cross, the recessive mutant ('MC') served as the maternal parents to ensure that self-pollen contamination which would be phenotypically evident. A total of 41 F2 plants surviving to the flowering stage were

scored for flower colour and *DFR* genotypes. For genotyping, genomic DNA was isolated from parental, F1 and F2 plants. *AN1* was sequenced from eight F2 plants with 'CP' and 'MC' half each, to test if F2 phenotypes correspond to the parental genotypes. Primers for DFR are F: 5'ACAGGGGCTTCAGGTTTCAT, R: 3'CTCAGGATCTTTGGATTCAA. Primers for *AN1* are e1F2/e6R1 given in **Table 4.2.**

DFR in the parental plants were sequenced. The genetic marker of a point mutation T/G at position 148 in exon 1 was used. The mutation interferes with the recognition site of the restrictive endonuclease XapI. The sequence in 'CP' (5'AAATTT) will be digested whereas 'MC' (5'AAATTG) will not.

4.3 Results

4.3.1 Pigment profiles

The white cultivars of *L. odoratus* and *L. sativus,* lacked visible nectar guides, and as expected contained no anthocyanidins (**Figures 4.2a**, **4.3a**, **4.3b**). Both had kaempferol (monohydroxylated at B ring) and quercetin (dihydroxylated) as the major flavonols. The white cultivars had more flavonols with the total flavonol ratio between the white and blue cultivar being 1.4 for *L. odoratus* and 2.7 for *L. sativus*.

The main anthocyanidins in *L. odoratus* 'CP' are delphinidin and its methylated derivatives petunidin and malvidin (**Figure 4.2**). The banner is dark reddish-purple, and contained more delphinidin and petunidin (3'-methylated) than the blue wings. In comparison, malvidin (3'5'-methylated) and peonidin (3' methylated derivative of the dihydroxylated cyanidin) were the main anthocyanidins in the blue *L. sativus* (**Figure 4.3**). The main flavonol aglycones in both *L. odoratus* and *L. sativus* were kaempferol and quercetin, with a trace amount of myricetin and isorhamnetin (methylated form of quercetin) only detected in *L. odoratus*.

Observing species of wild origin, the main anthocyanidins in *L. clymenum* were the delphinidin series.

Several white or yellow forms had small amounts of anthocyanin-like pigment, either in the dark nectar guides, or developing as floral colour changed during aging (**Table 4.3**). However, malvidin was only detected in one of the species *L. chrysanthus*, and no other albescent or yellow species had detectable anthocyanidins (**Figure 4.4a**). The flavonol aglycone profile was diverse across species of wild origins: the yellow *L. annuus* had quercetin, the white *L. ochroleucus* had kaempferol and quercetin, the cream *L. ochrus* had kaempferol, quercetin and isorhamnetin, the yellow sister species *L. chrysanthus* and L. *chloranthus* had all of the above plus trace amount of myricetin (Figure 4a). The yellow in *L. chrysanthus* is intense and corresponded to a much higher flavonol content than *L. chloranthus*. In fact, out of the six 'albinos', *L. chrysanthus* had the most flavonols. And finally the yellow *L. aphaca* also had all of the surveyed flavonols with the trihydroxylated myricetin being the most abundant. In addition, myricetin was either in trace amounts or could not be detected at all in the other white or yellow species. It was however, present relatively strongly in the red *L. clymenum*. Overall, a cream/white flower did not have lower level of flavonol than the yellow ones.

For the analysis of separate floral parts, flavonol abundance was similar in wings and banner petals for both *L. aphaca* and *L. chrysanthus*, except for quercetin which varied more between floral parts. The wings of *L. chloranthus* had more flavonol than the banner even though the flower is concolourous. *L. chrysanthus* also had more flavonol in the wings.

4.3.2 Trans-regulatory silencing of *DFR* is associated with a white mutant of *L. odoratus*

Genetic crosses between 'CP' and 'MC' indicated that the colour trait segregated as a single locus, with purple dominant to white (**Table 4.5**). As DFR is the first enzyme committed to anthocyanin production, I tested its expression in the white cultivar 'MC'. *L. odoratus DFR* is 939 bp or 342 amino acids long. A blast search matched it with *DFR* from *Pisum* (97% sequence identity), *Medicago* (91% identity) and *Lotus* (89% identity). In the RT-PCR analysis, 'MC' did not express *DFR* (**Figure 4.2b**). It is also the only structural gene tested that was not expressed in the white mutant. Although structural genes are often co-regulated in the anthocyanin biosynthetic pathway (Mol et al., 1998), in 'MC' both *anthocyanin synthase* (*ANS*) and *flavonol synthase* (*FLS*, results not shown) were expressed, indicating in the sweet peas, *DFR* was regulated separately from the downstream enzymes.

Out of the five TFs tested for expression, *AN1* and *ROI1* (*ROSE INTENSITY1*) homologues were not expressed in 'MC'. Those TFs are from the anthocyanin regulatory network and are known to control *DFR* expression in systems such as *Petunia, Antirrhinum, Aquilegia* and *Arabidopsis*. Anthocyanin regulatory genes belong to four gene families: *bHLH (AN1*), R2-R3 MYB (*AN2, AN4*) (Quattrocchio et al., 1999), single repeat R3 MYB (*ROI*) (Yuan et al., 2013) and WD40 (*AN11*) (Spelt et al., 2000). Both *bHLH* and *R3 MYB* was thus involved in *DFR* silencing in the white sweet pea.

In 'CP' where banner and wings were analyzed separately, the expression levels of the structural gene *DFR* and *ANS* seem to be lower in the banner than in the wings of 'CP' whereas most TFs show similar level of expression, except for *ROI* which is lower in the wings.

The co-segregation analysis showed that among the 41 surviving F2 plants, all 12 white individuals lacked *DFR, AN1 and ROI* expression. However, *DFR* did not co-segregate with the white phenotype implying that *DFR* expression is controlled by a separate locus (**Table 4.5**). The TF *AN1* also did not co-segregate with the white phenotype either, suggesting a *trans*-regulatory control of the *AN1*. *ROI* was not included in the co-segregation analysis as no genetic marker was found.

4.3.3 Lack of *DFR* expression in a white cultivar of *L. sativus*

As in *L. odoratus*, *DFR* was not expressed in the semi-domesticated *L. sativus* var. *albus.* However, unlike the white *L. odoratus*, *AN1* was expressed in the white *L. sativus*, indicating a different regulatory mechanism leading to *DFR* silencing in the latter.

4.3.4 *DFR* expression in *Lathyrus* of wild origins

Unlike the domesticated cultivars, none of the yellow or cream *Lathyrus* of wild origin showed complete silencing of *DFR.* While *DFR* expression was low in the two white species, it varied from very low to very high among the yellow species. In particular, the yellow *L. chloranthus* (no visible nectar guide) had a comparably high expression to the blue *L. hirsutus*, and is higher than the red *L. tingitanus*. Another yellow species, *L. chrysanthus*, had a very low expression, despite its black nectar

guide.

4.4 Discussion

4.4.1 Molecular similarity, but not genetic parallelism, in phenotypically parallel event

Competing with *FLS* for common substrates (dihydroflavonols), *DFR* is the first enzyme in the anthocyanin pathway that commits the flux to anthocyanin production. Mutations affecting *DFR* function have frequently been associated with white floral phenotypes, usually via a *trans*-regulatory mechanism (Streisfeld and Rausher, 2009, Whittall et al., 2006) although colour loss due to *cis* mutations at *DFR* have been reported. Those *cis* mutations can either cause a loss of function in *DFR* such as in soybean (Yan et al., 2014) or have no substantial effect on *DFR* expression as in *Mimulus lewisii* (Wu et al., 2013).

Both albino cultivars of *L. odoratus* and *L. sativus* lacked *DFR* expression. However, without a functional assay, causality is not established. *DFR* silencing inactivated the pathway branch for anthocyanin production. In *L. odoratus*, the unidentified white factor is a *trans* mutation to *DFR* and *AN1*, and is either a *cis* or *trans* to *ROI*. This white factor in *L. odoratus* is different from that in the classical system *Petunia*, where *AN1* expression is regulated by *AN2* and *AN4* (Spelt et al., 2000) whereas in *L. odoratus*, neither *AN2* nor *AN4* homologue co-segregated with the white phenotype. In another system *Mimulus*, the recessive allele of an MYB factor *ROI1* makes more concentrated anthocyanins, and a dominant mutation is responsible for the transition from a red to a pink species. The lack of *ROI* expression in the white *L. odoratus* suggests that it may have a novel function. As no genetic marker was found, a *cis* mutation at *ROI* cannot be ruled out as the white factor. However, I was restricted to studying only one of the two white factors in *L. odoratus*. I attempted to find a different genetic strain of white by crossing 'MC' with another white cultivar 'Dorothy Eckford' but the resulting hybrid was also white, indicating no complementation.

For *L. sativus*, without a co-segregation analysis, it is not clear whether the mutation is at a locus *cis* or *trans* to *DFR*. But even in the case of a *trans*-regulation, the regulatory route is not via *AN1* down-regulation, hence it is likely to use a different transcription factor from the *L. odoratus* case. In a survey

comparing gene usage in parallel and convergent evolution among natural populations, the probability of using the same gene declined as taxa get more phylogenetically distant (Conte et al., 2012). The two white mutants came about under artificial selection through similar genetic mechanisms involving down-regulation of the same gene *DFR*, but different regulatory genes are likely to be the underlying causes.

In addition, other structural genes of the anthocyanin pathway are often co-regulated with *DFR*. For instance, in *Ipomoea alba DFR* was co-regulated with *CHS, CHI* and *DFR* (Clegg and Durbin, 2003, Durbin et al., 2003), whereas in the white *Aquilegia DFR* was co-regulated with *ANS* (Whittall et al., 2006) and in the yellow *Mimulus DFR* was down-regulated together with *F3H* and *ANS* (Streisfeld and Rauscher, 2009). I only tested the expression of *FLS* and *ANS* in 'MC', and did not find any down-regulation. In future work, a transcriptomic study could be used to reveal the regulatory network in the white sweet pea, in a more comprehensive manner. However, if I were to adopt the candidate gene approach, anthocyanin pathway genes *CHS*, *F3H* and *UFGT* are candidates worthy of further investigation.

4.4.2 Varied *DFR* expression in yellow and white wild species

None of the *Lathyrus* of wild origin showed complete *DFR* silencing, indicating other genetic route(s) have contributed to the major loss of colour. A few genetic possibilities that may lead to pigment loss, without a complete *DFR* silencing include: **1)** another structural gene, instead of *DFR*, is completely silenced, **2)** a structural mutation in *DFR* affects the protein function without an apparent change in expression, **3)** another *DFR* copy is used and **4)** down-regulation of other structural genes (result in lowered expressions) together with *DFR* leads to a lack of the final products since intermediate substrates are scanty.

Scenarios **1** and **4** are most plausible and could be tested using a transcriptomic approach. Scenario **2** is unlikely since a mutation affecting DFR protein function will have deleterious consequences due to functional pleiotropy in physiological factors such as stress-response, which is required in the protection against UV irradiation, oxidative and heat shock stress (Lorenc-Kukula et al., 2005).

Regarding scenario **3**, *DFR* is a multi-copy gene although only a single copy is responsible for floral colour formation in many plants such as *Petunia* (Meyer et al., 1987), *Antirrhinum* (Almeida et al., 1989), *Viola* (Farzad et al., 2003), *Ipomea* (Inagaki et al., 1999) and *Lotus* (Ojeda et al., 2013). At least two *DFR* copies exist in *Lathyrus*, as determined from blasting the *Lathyrus* transcriptome library against a *Petunia* library. But only one of them was functional in *Lathyrus odoratus*. However, multiple functional copies are possible, especially with recent polyploidy events which result in gene duplication after species divergence. *Lathyrus* by and large is a diploid (2n=14) genus with some polyploid species in perennial *Lathyrus* such as the hexaploid *L. palustris* and the tetraploid *L. venosus* (Wylie and Darlington, 1955). Since none of the species tested here is a polyploid, it is unlikely they would use a different *DFR* copy than *L. odoratus.* However, this scenario **3** is still possible for another yellow species *L. pratensis* which has a tetraploid variant (Gutierrez et al., 1994).

Under natural selection, *DFR* expression in the largely unpigmented species was probably maintained in some cases for nectar guide colouration and in others for producing stress-induced anthocyanins which act as antioxidant. In *L. chrysanthus*, anthocyanidin was found exclusively in the banner (with black nectar guide) suggesting a nectar guide origin of the pigment. A low *DFR* expression (**Figure 4.4b**) matched with this low amount of malvidin in the banner. This nectar-guide-specific expression of *DFR* should be apparent if the veins were separately analyzed. The low level of *DFR* expression could be due to either dilution from a non *DFR*-expressing petal background or the use of another *DFR* copy at the vein. Previous study shows that duplicated *DFR* genes lead to novel function at the petal spots in *Clarkia* (Martins et al., 2013). However, the *L. chrysanthus* case is different from that of *Clarkia:* the temporal and compositional variation of pigment production in the *Clarkia* spot (malvidin, appears early) and the pink petal background (cyanidin/peonidin, appears late) indicates that neofunctionalization of the duplicated *DFR* is a necessary step to yield a different pigment product at different time, from those of the ancestral *DFR*. Hence, low *DFR* expression in *L. chrysanthus* is likely due to a dilution effect, assuming, as is likely, that *L. chrysanthus* uses the same *DFR* copy as *L. odoratus* and *L. sativus*.

However, in the case of *L. ochroleucus,* a white species without any nectar guide, *DFR* expression

may be explained by its role in stress response. *L. ochroleucus* turns yellowish brown with age, presumably those pigments act as antioxidant during senescence against oxidative deterioration (Cavaiuolo et al., 2013). But the pigment contributing to this colour change is not known so it is not clear if *DFR* is involved. Other species such as *L. chloranthus* and *L. annuus* turn brown and red respectively in response to heat or excess sunlight in the greenhouse (personal obs.). So it is likely that *DFR* is required for protection against oxidative stress. Both UV-B and UV-A have been suggested to induce anthocyanin accumulation in plants (Gould, 2004, Guo et al., 2008, Morales et al., 2010). Mutant plants that lack epidermal flavonoids showed a UV-induced injury (Landry et al., 1995). In a meta-analysis of 450 field reports, it is concluded that in response to an elevated UV-B (280-320 nm), plants increase the accumulation of UV-absorbing compounds (Searles et al., 2001). Moreover, total phenolics in *Pisum sativum* increased with an elevated UV-B level (Hatcher and Paul, 1994). On a molecular level, in *Arabidopsis*, UV-B induces the expression of *DFR* together with *PAL*, *CHS* and *CHI* (reviewed in (Jenkins et al., 1997). However, it is also argued that anthocyanin accumulation is a by-product of a general stress response involving other plant hormones, and therefore is non-adaptive (Loreti et al., 2008), although this hypothesis has not been tested.

In addition, it seems that domestication is permissive to major phenotypic change compared to the wild progenitors, and domesticates can have morphological modifications that are too extreme to survive in the wild.

4.4.3 Ecological context of the 'white shifts'

Shifts in floral colour can be due to either adaptive or non-adaptive mechanisms (Rausher, 2008). In a few other species the 'white shifts', together with other change in floral features (pollination syndrome) is associated with a pollinator switch. Classical example includes *Petunia* and *Aquilegia*. The magenta *Petunia integrifolia* is pollinated by bees whereas white *P. axillaris* with a loss-of-function in *AN2* is mainly pollinated by nocturnal hawk moth (Hoballah et al., 2007a). *Aquilegia formosa* (red) is pollinated by hummingbirds whereas *A. pubescens* (white) is pollinated by hawkmoths (Fulton and Hodges, 1999).

In *Lathyrus*, five members of the series *Lutei* (*L. luteus, L. linnaei, L. emodii, L. gmelinii and L. laevigatus*) sharing synapomorphies such as yellow, dense, small racemes (Bässler, 1973) are possible to result from a pollination syndrome change and the yellow shift from their coloured relatives might be adaptive. However, except one member *L. luteus*, which was recorded to be visted by bumble bees (Knuth et al., 1906), no field observation of pollination is available for other members of this group, therefore it is not possible to confirm this.

In *Antirhinum*, it has been shown that white flowers are generally visited less often (Glover and Martin, 1998), but field observations or experimental data are required to confirm the applicability of this observation to *Lathyrus*. Where loss of anthocyanin and transition to white or yellow flowers has occurred in *Lathyrus* without any apparent other changes of features, colour loss might be either adaptive to non-pollinator agents or non-adaptive. I have observed bees visiting both blue and white cultivars indiscriminately. Future studies in *L. odoratus* and *L. sativus* could include measuring nectar amount and concentration, as well as pollinator visitation rate.

Most white or cream *Lathyrus* have flavonol or flavone in the petal. Why do the white flowers not completely lack any pigment? Two explanations are plausible: **1)** the glycosylated form of flavonol and flavone absorb strongly at the UV region hence can provide nectar guides. Pure white flower without nectar guide is rare in nature. This natural selection against albinism may be a result of the inferior nectar guides (reduced contrasts) and the subsequent longer handling time due to partial discrimination by optimally foraging pollinators (Waser and Price, 1983). Surprisingly, the functional role of nectar guide is not very well-documented in literature, but the visual line or dot shape forming a contrast of floral parts is presumably important for the close-up orientation for the optimally-foraging pollinators. This contrast can happen in the UV region, making it indistinct to humans. I only tested two white species, and one of them (*L. ochrus*) has visible nectar guide. The other white species *L. ochroleucus* has no visible nectar guide but the presence of flavonols in the petal may be explained by **2)** UV-A and UV-B irradiation induces free radical, superoxide and H_2O_2 in plants which causes oxidative stress. It is widely accepted that phenylpropanoid compounds are synthesized as an adaptive response to UV exposure, to act as ‘sunscreens’ since they can scavenge free radicals (Caldwell et al.,

2007).

Overall, colour transition in *Lathyrus* might represent a mixed case of adaptive and non-adaptive processes, since there is no obvious evidence of major pollination changes, such as shift of pollinator accompanying colour change. In future I could include a survey of colour polymorphisms within species to see how often a loss of colour would happen intra-specifically, and to assess if any correlation between an array of ecological factors (both biotic and abiotic) and the loss of colour trait exist. But as of now, loss of anthocyanin colour, same as many other phenotypes of wild species, could be the result of natural selection, drift and phylogenetic constraint, or a mix of all three (Rausher, 2008).

Table 4.1 Regulators of the anthocyanin pathway in four model plants (*Petunia, Antirrhinum, Arabidopsis* and *Maize*).

Family	*Petunia*	*Antirrhinum*	*Arabidopsis*	Maize
MYB	*AN2, AN4* (Quattrocchio et al., 1999)	*VENOSA/ROSEA (Schwinn et al., 2006)*	*TT3* (MYB75)	*A1/PL1* (Paz-Ares et al., 1987, Cone et al., 1993)
bHLH	*AN1* (Spelt et al., 2002)	*DELILA/MUTALIBIS* (Goodrich et al., 1992)	*TT8* (Nesi et al., 2000)	*R (Ludwig and Wessler, 1990)*
WD40	*AN11* (de Vetten et al., 1997)		*TTG1* (Zhang et al., 2003)	*PAC1* (Carey et al., 2004)

Table 4.2 Primers used in RT-PCR. Genes tested are the *Lathyrus* homologues of the structural genes *DFR2* (expressed in *L. odoratus*) and *ANS*, and transcription factors *AN1*, *RO1*, *AN2*, *WD40*.

Taxa	Gene	Primer	Sequence (5' – 3')
All taxa	*ACTIN2*	F2 R1	CCGGTTCTTCTCACTGAAGC GGCTGGAACAGAACTTCAGG
L. chrysanthus *L. sativus* *L. sativus* var. *albus*	*DFR2*	F5+6 R1	GGAAATGAAGCACATTACAC AGCCCTTTTTCTCTGCAGGT
L. annuus *L. chloranthus* *L. hirsutus* *L. ochrus* *L. ochroleucus* *L. odoratus* *L. tingitanus*	*DFR2*	F1 R1+2	ACAGGGGCTTCAGGTTTCAT CTCAGGATCTTTGGATTCAA
L. odoratus	*ANS*	F1 R1	AAAGAGACAGTGCTCAAGAAC GGTCTGAATGAGAAGACAGC
L. sativus *L. chrysanthus* *L. aphaca*	*AN1*	e1f1 e6R1	GCCCACAACAATTGATACTGG GTGTGTGTCTTCTTGCGTTAA
L. odoratus	*AN1*	e1F2 e6R1	TGCTGAAGAGGCTTCTCTG GTGTGTGTCTTCTTGCGTTAA
L. odoratus	*RO1*	F3S R2b	ATGGCCAACGTGGAACACAC ATTGGCTAGTGGAAGAGTCTC
L. odoratus	*AN2*	F2 R1	GAGAAAATCCGAGGATGTGG AGACAATATGGAGAAGGGACAT
L. odoratus	*WD40*	F1 R1	ACTTCTTTTGATTGGAATGANG GGAGCCCAAGCAATGGCATT

Table 4.3 Plant source, colour and relative expression levels for *DFR*. The presence or absence of pigmented nectar guides (radiating lines), and colour-changing during aging are also indicated for the unpigmented taxa. Sources: (RG), Renee's Garden; (RP), Roger Parsons; accession and origin were indicated for USDA germplasm seeds. Gene expression levels for *DFR* are scored as: -, not detected; +, ++, +++, low, moderate and high expression, according to qRT-PCR results (see **Figure 4.2-4.4**).

Taxa	**Source and origin**	**Colour**	***DFR***
L. odoratus cv. 'Cupani'	Commercial (RG)	Purple	+++
L. odoratus cv. 'Mrs Collier'	Commercial (RG)	White (no guide)	-
L. sativus	Commercial (RG)	Blue	+++
L. sativus var. *albus*	Commercial (RP)	White (no guide)	-
L. clymenum	PI 283503, Portugal	Purple	n/a
L. hirsutus	PI 283521, Tunisia	Purple	+++
L. tingitanus	Commercial (RP)	Magenta	++
L. aphaca	PI 286528, Afghanistan	Yellow (black guides)	n/a
L. annuus	Commercial (RP)	Yellow (red guides, turns red)	+++
L. chloranthus	PI 229794, Iran	Yellow (no guide, turns brown)	++
L. chrysanthus	PI 604496, Syria	Yellow (black guides)	+
L. ochroleucus	Wild collected near Kamloops, BC, CA	Cream (no line, turns yellow)	+
L. ochrus	Commercial (RP)	Cream (brown lines)	+

Table 4.4 Peak assignment for the analysis of the aqueous methanol extract of the flowers of several *Lathyrus* species, based on the flavonoid aglycones standards. T_R (retention time) decreases with higher polarity of the molecules, which elute faster in a reverse phase LC. m/z (mass to charge ratio) calculation is based on the mass database at massbank.jp. +/- indicates the mode the compound is detected with. Absorption maxima followed manufacturer's certificate of analysis.

Pigment	**t_R (min)**	**m/z**	**UV λ_{max} (nm)**
Cyanidin	1.4	287(+)	538
Peonidin	2.8	301(+)	537
Delphinidin	0.9	303(+)	548
Malvidin	3.1	331(+)	547
Petunidin	1.8	317(+)	547
Quercetin	4.5	301(-)	372
Kaempferol	5.8	285(-)	365
Myricetin	2.8	317(-)	374
Isorhamnetin	5.9	315(-)	372

Table 4.5 F2 co-segregation analyzed with respect to the colour phenotype and *DFR* genotype in 41 F2 individuals obtained from the cross 'CP' (purple) × 'MC' (white). White is the recessive phenotype (purple: white = 3:1) and ***b*** ('MC' *DFR* marker) is the recessive genotype (BB:Bb:bb 1:2:1, ***B*** is the 'CP' *DFR* allele). The white phenotype has an underlying recessive genetic factor associated with it. But if this locus and *DFR* were of the same nature, a 'bb' genotype would have always corresponded to a 'white' phenotype, which was not observed.

Individual	**Genotype**	**Phenotype**	**Individual**	**genotype**	**Phenotype**
1	Bb	**White**	22	Bb	Purple
2	**bb**	Purple	23	Bb	Purple
3	BB	**White**	25	**bb**	Purple
4	**bb**	Purple	26	BB	**White**
5	Bb	**White**	27	Bb	**White**
6	Bb	Purple	28	BB	Purple
7	BB	Purple	29	**bb**	Purple
8	Bb	Purple	30	Bb	Purple
9	Bb	Purple	31	Bb	Purple
10	BB	Purple	32	Bb	Purple
11	BB	**White**	33	BB	Purple
12	Bb	Purple	34	Bb	**White**
13	**bb**	**White**	35	BB	Purple
14	**bb**	Purple	36	Bb	Purple
15	Bb	Purple	37	BB	**White**
16	BB	**White**	38	**bb**	**White**
17	Bb	Purple	39	Bb	Purple
18	**bb**	Purple	40	Bb	Purple
19	**bb**	Purple	41	Bb	**White**
21	**bb**	Purple	42	BB	Purple
			43	**bb**	Purple

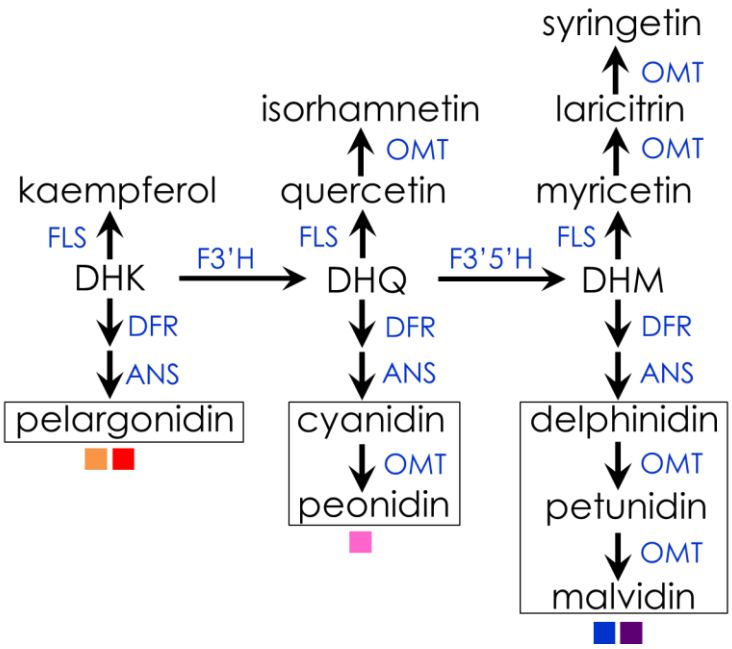

Figure 4.1 Partial scheme of the anthocyanin biosynthetic pathway. Dihydroflavonol DHK (dihydrokaepmferol), DHQ (dihydroquercetin) and DHM (dihydromyricetin) can be converted to either the flavonols, catalyzed by flavonol synthase (FLS), or to anthocyanidins, catalyzed by dihydroflavonol reductase (DFR) and anthocyanin synthase (ANS). Methylation of anthocyanins and flavonols are carried out by *O*-methyltransferase (OMT). The colour of the end products of each pathway is indicated, delphinidin derivatives are associated with blueness, cyanidin derivatives are pink or reddish and pelargonidin gives a characteristic bright reddish orange colour. Flavonols differ in the number and type of substitution in the B ring. Kaempferol is monohydroxylated at the C3 position, quercetin is dihydroxylated at C3 and C4 positions and myricetin is trihydroxylated at C3, C4 and C5. Isorhamnetin is the methylated form of quercetin, and laricitrin and syringetin are the methylated forms of myricetin.

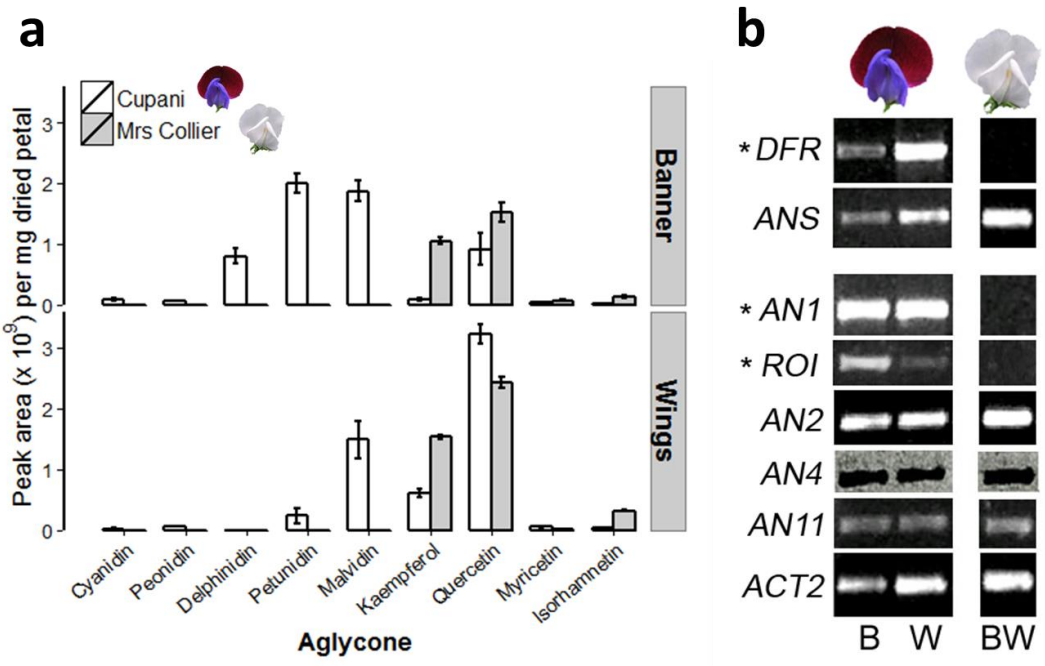

Figure 4.2 (**a) Pigment profiles measured with HPLC-MC, compared for purple and white cultivars of *L. odoratus***. **(b) RT-PCR results for anthocyanin pathway structural genes (*DFR, ANS*) and five regulatory genes**. Regulators belong to various gene families: MYB (*AN2, ROI, AN4* homologues), bHLH (*AN1* homologue) and WD40 repeat (*AN11* homologue). Petal parts designation: B, banner; w, wings.

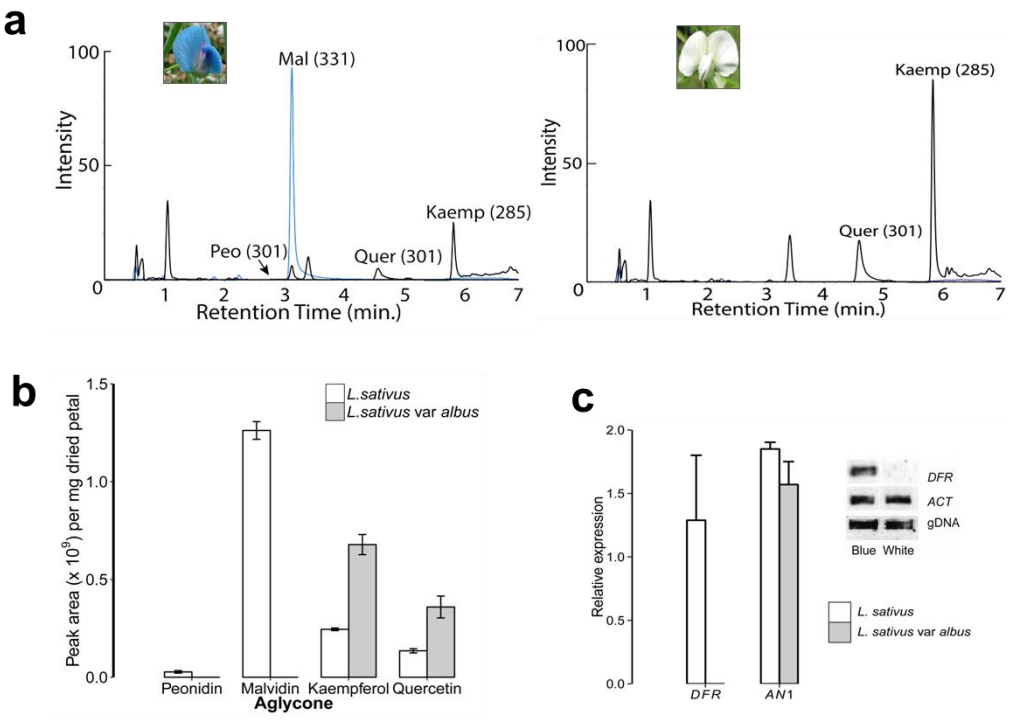

Figure 4.3 Pigment profile and expression of *DFR* in another domesticated species *L. sativus* and the white variety *L. sativus* var. *albus*. (a) LC chromatograms. **(b)** Pigment abundance graph and **(c)** RT-PCR results for *DFR* and its regulator *AN1*, which is co-silenced in the sweet pea example.

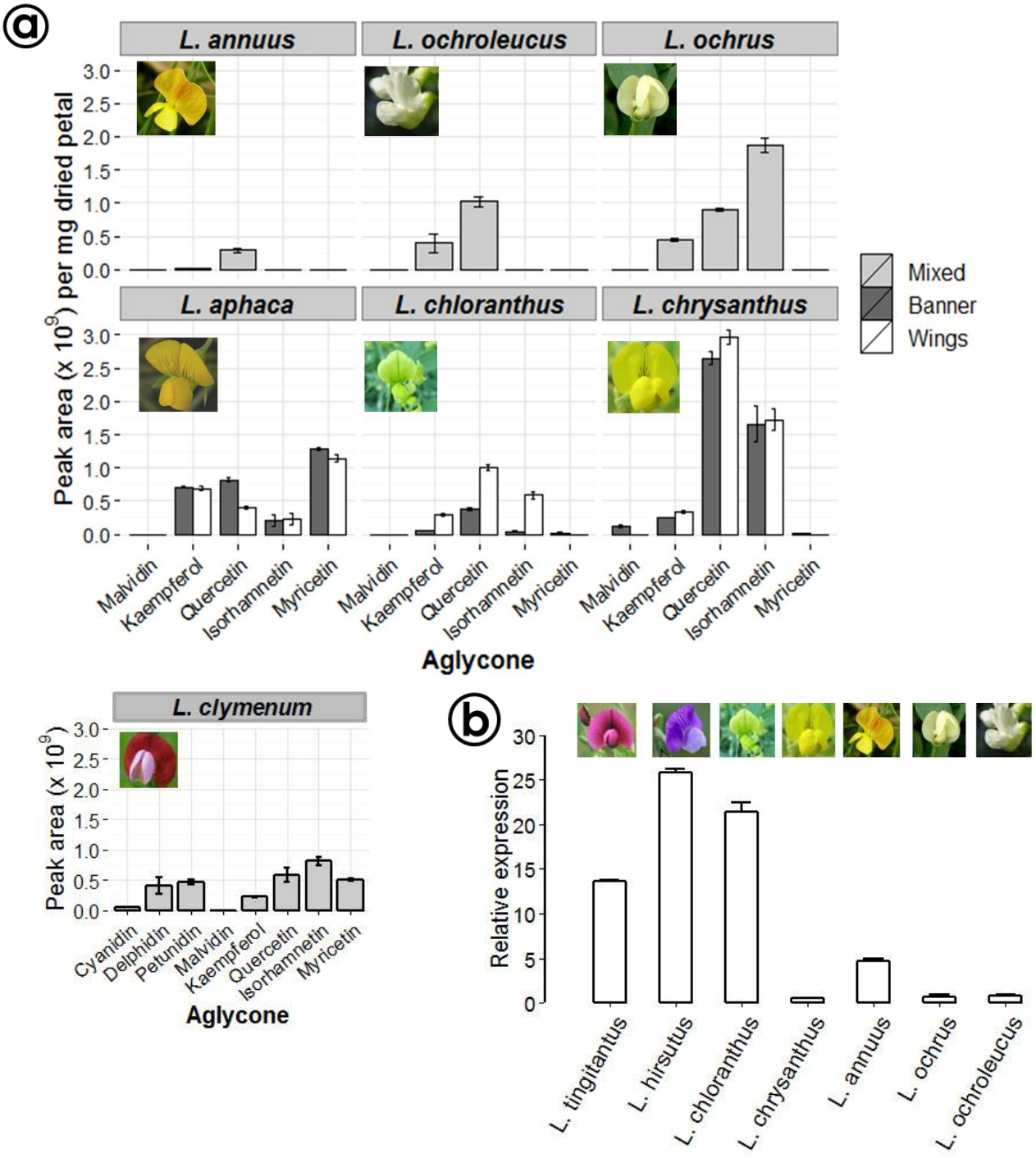

Figure 4.4 Pigment profile and *DFR* expression. (a) Relative abundance of flavonoid aglycones in six unpigmented (upper two panel) and one coloured species (lower left) of *Lathyrus.* Water-soluble pigments were extracted from silica-dried petal, harvested at anthesis. Samples were analyzed with LC-MS. The total area under an MS peak was estimated as the relative abundance for that flavonol or anthocyanin aglycone. **(b) Quantitative real-time RT-PCR of *DFR* expression in pigmented versus unpigmented *Lathyrus*.**

Chapter 5. Conclusions and Future Directions

The major aim of this thesis was to study the flower colour evolution in the colour-rich genus *Lathyrus* (containing 150~160 species). Overall, both structural and regulatory mutations are involved in the colour shifts in *Lathyrus*. Work presented in this thesis is an extension of early studies of Mendelian genetics in the sweet pea (*L. odoratus*) (Bateson et al., 1906, Bateson et al., 1909, Fleming, 1925, Beale et al., 1939) while placing it in a molecular and an evolutionary context. The ancestral reconstruction of colour-related traits used a multi-species coalescent tree that was constructed using DNA sequence data mostly retrieved from previous phylogenetic studies (Kenicer et al., 2005, Schaefer et al., 2012).

The genus *Lathyrus* occupies a wide range of biogeographic regions from temperate northern hemisphere to tropical East Africa and Andean South America. Most members of the genus are mesophytes that grow in open woodlands, forest margins and road verges with littoral, alpine and drought-tolerant species also represented (Kenicer et al., 2005). Such habitat diversity provides opportunities for ecological adaptations involving floral colour, when pollinators or abiotic agents (such as heat stress and UV radiation) may impose natural selection either directly or indirectly via pleiotropic effects; or when selection favours certain vegetative traits determined by loci linked to floral ones. Non-adaptive processes (i.e. drift) may also contribute to floral colour diversity in *Lathyrus* but these are not within the scope of this thesis. The work presented here should open up new avenues for further investigation of this genus. *Lathyrus* is largely unexplored in terms of evolutionary-developmental and molecular ecological studies.

5.1 Summary

5.1.1 Chapter 2

I discovered that the historical *A1* locus of *L. odoratus* corresponds to a missense mutation (G/A) in the first exon of *F3'5'H* in red and white bicoloured 'PL'. The resulting amino acid change, from glycine to aspartic acid, happened at a highly conserved site of the substrate recognition site 1 (SRS1). However, heterologous transformation of *F3'5'H* into *PAP1D*, an *Arabidopsis* line that overexpresses

anthocyanins, did not yield the delphinidin series as expected. Instead, both *F3'5'H*-transgenic lines produced much higher cyanidin, which is at a drastically higher level in the mutant *plF3'5'H* transgenics than that in Columbia and *cpF3'5'H* transgenics. What caused this? All F3'5'H enzymes have some F3'H activity (Seitz et al., 2006). I speculate the mutant enzyme is toggled from primary F3'5'H activity to a relatively efficient F3'H (with little F3'5'H activity), through unknown enzyme kinetics. Due to the absence of F3'5'H, plants such as *Rosa* and *Gerbera* are not able to synthesize the blue delphinidin pigments. But the blue to red/pink transition is not always caused by *F3'5'H*, such as in *Ipomoea purpurea* where the red flower variant is caused by transpositional insertion leading to a premature stop codon in *F3'H* (Zufall and Rausher, 2003). The finding of a single nucleotide change contributing to the blue to pink transition in the sweet pea adds to the current consensus of blue colour–F3'5'H association and, even more importantly, to the rare cases of a single nucleotide mutation contributing to drastic morphological change in nature.

A1 is the fifth locus in *L. odoratus* that has been identified, following the flowering time locus *SP*, branching locus *B* (Ross and Murfet, 1988), the tendril-less locus *T* (Hofer et al., 2009), and the flower symmetry locus *A3* (Woollacott, 2010). The finding has potential evolutionary implications because a red variant equivalent to the 'PL' phenotype was reported in the wild (Bateson, 1905), although no contemporary record can confirm this.

5.1.2 Chapter 3

A Bayesian multi-species coalescent tree was built for 92 *Lathyrus*, two *Pisum* and 11 outgroups using BEAST. *Lathyrus-Pisum* forms a monophyletic clade, agreeing with previous studies (Kenicer et al., 2005, Schaefer et al., 2012). The ancestral states for *Lathyrus* life history and flower traits were inferred to be annual, anthocyanin rich (pink to purple) and concolourous, using ML reconstruction. Loss of anthocyanin colour happened at least 16 times with no reversal. Transition from annual to perennial happened at least four times with three reversals. When including the white species, transition from concolourous to bicolourous pattern occurred at least 16 times with five reversals.

Unlike the life cycle and floral pattern traits, loss of floral colour in *Lathyrus* is asymmetrical (or

unidirectional) with no reversal inferred. This agrees with the generalization that changes from blue species to red, yellow or white species happen more often than the reverse, as observed in many other angiosperm groups such as *Ipomoea* (Rausher, 2008), *Penstemon* (Wilson et al., 2004), *Iochroma* (Smith et al., 2013) and *Sinningieae* (Perret et al., 2003). This is mainly because a loss-of-function mutation is generally irreversible and such irreversibility can be reinforced if two or more mutations are required for a transition (Rausher, 2006).

In addition, both colour and life history traits showed strong phylogenetic signals indicated by Pagel's lambda values. And no correlated evolution was found between the two traits when phylogenetic independent contrast was taken into consideration. It was different from expectation as perennial *Lathyrus* are strictly outcrossing (Brahim et al., 2001) and therefore rely on pollinators so they are required to have a more attractive floral display. Annuals, on the other hand, are capable of selfing which is often associated with a less attractive floral display. However, the lack of association could be due to the fact that the correlation analysis is limited since all members (42 taxa) of the section *Orobus* are perennial and the majority of *Lathyrus* are pigmented.

5.1.3 Chapter 4

I found that *DFR* expression was associated with the loss-of-colour in the sweet pea white mutant 'MC'. The sweet pea orthologues of *AN2* (MYB) and *AN1* (bHLH) were also not expressed. But in the F2 co-segregation analysis, none of the silenced gene loci were linked with the white phenotype. So the corresponding gene for the white locus (*R*) was not found. *DFR* expression was also abolished in the white mutant of a domesticated crop – the grass pea (*L. sativus*). But *AN1* was expressed in the albino grass pea, indicating a different regulation mechanism. Furthermore, *DFR* was never completely silenced in the yellow and cream wild *Lathyrus* species that had originated and been maintained under natural rather than artificial selection, even though no anthocyanin could be detected in most of the species tested. *DFR* expression in the cream species is in general lower than that in the yellow ones. I conclude that repeated morphological evolution is accompanied by similar changes at the biochemical level, but the molecular mechanisms are less consistent, likely because of pleiotropic effects.

Anthocyanin synthesis requires the expression of several TFs (*MYB, bHLH* and *WD40*) and structural genes (key genes include *CHS, DFR, ANS*). Many of the TFs in floral colour transition were first identified in a few plants such as *Petunia* (Wiering, 1974, de Vetten et al., 1997, Quattrocchio et al., 1999, Spelt et al., 2000, Hoballah et al., 2007a) and *Antirrhinum* (Almeida et al., 1989, Comba et al., 2000, Schwinn et al., 2001, Schwinn et al., 2006). The transcription factor corresponding to the *R* locus in *Lathyrus* is different from all of the previous findings in other systems.

5.2 Future directions

For the functional verification of *F3'5'H* in 'PL', multiple insertion lines should be included in the heterologous transformation of sweet pea *F3'5'H* into *Arabidopsis* to eliminate positional bias. For the genetic basis for the loss of colour trait: a transcriptomic approach could be applied to the white mutants and unpigmented *Lathyrus* of wild origin. For instance, tissues of the nectar guides could be separated from the rest of the corolla and these test for any differential *DFR* expression.

In addition, within a species, I only studied white mutants which are the product of artificial selection. To test the effect of artificial versus natural selection on gene usage below the species level, one can test the *DFR* expression in naturally occurring white morphs of a coloured taxon. For example, *L. nevadensis* has white morphs in the southern range of the population (Broich, 2007). An interesting question would be: Does the white *L. nevadensis* express *DFR*?

It will be extremely beneficial to apply the knowledge gained here in an ecological context. Possible directions include finding out if a red/white bicoloured variant of the sweet pea still exists in its native habitat, Sicily, and then to infer if standing genetic variation contributes to evolution (for instance through balancing selection). Pollinator visitation experiments comparing the pigmented perennials versus yellow/cream perennials in their natural habitats could answer the question of whether colour confers fitness differences and is under selection.

5.3 Final remarks

The work presented here on floral colour evolution in the genus *Lathyrus* touches on a few central

questions in evolutionary biology concerning genetic architecture and the genetic basis of traits.

1) The nature of genes involved in adaptation.

Growing evidence indicates that *cis*- or *trans*-regulatory machinery contribute more to morphological evolution whereas structural genes contribute predominantly to physiological adaptations (reviewed in Durbin et al., 2003). In this study, floral colour change, a morphological evolutionary trait, is variously controlled by mutation at a structural gene (the blue to pink transition) and a mutation at a *trans*-regulatory gene (blue to white transition) in *L. odoratus*. This clearly shows that within one single species, the same type of morphological change can be achieved through both ways and suggests that generalizations will be difficult to make.

2) Evolution of traits determined by few genes of large effect, or many genes of small effect.

In the sweet pea, colour mutants maintained under selective breeding are controlled by a single gene, fitting the first scenario. In comparison, cream and yellow *Lathyrus,* originating and maintained in a natural environment, may be due to the latter scenario.

3) Role of chance variation in floral diversity.

Conspicuous flowers evolved to facilitate pollination. However, despite the important role of pollinator selection on directing floral features including colour (e.g. Muchhala et al., 2014), it is certainly not the sole factor shaping the current floral diversity on earth. It is entirely possible that sequentially appearing chance variation may account for at least some of the endless diversity of floral morphologies (Beatty, 2010). Such generation of variation itself may give direction to evolution and is the ultimate source of evolutionary creativity (Beatty, 2010). The environment can act on existing variation through selection but when such variation is neutral (neither advantageous nor disadvantageous to the survival of the individual), it can still be maintained and may be fixed by drift. I speculate that random drift is the dominant force shaping the diverse colour forms in *Lathyrus*.

Bibliography

Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell,* **26**: 962-80.

Almeida J, Carpenter R, Robbins TP, Martin C, Coen ES. 1989. Genetic interactions underlying flower color patterns in *Antirrhinum majus*. *Genes & Development,* **3**: 1758-1767.

Altshuler DL. 2003. Flower color, hummingbird pollination, and habitat irradiance in four neotropical forests. *Biotropica,* **35**: 344-355.

Arista M, Talavera M, Berjano R, Ortiz PL. 2013. Abiotic factors may explain the geographical distribution of flower colour morphs and the maintenance of colour polymorphism in the scarlet pimpernel. *Journal of Ecology,* **101**: 1613-1622.

Armbruster WS. 2002. Can indirect selection and genetic context contribute to trait diversification? A transition-probability study of blossom-color evolution in two genera. *Journal of Evolutionary Biology,* **15**: 468-486.

Asmussen C, Liston A. 1998. Chloroplast DNA characters, phylogeny, and classification of *Lathyrus* (Fabaceae). *Am J Bot,* **85**: 387.

Ayaz EMINE, Ertekin AS. 2008. Karyotype analysis of two species of genus Lathyrus from Southeastern Anatolia, Turkey. *Int. J. Agri. Biol,* **10**: 569-572.

Bak S, Beisson F, Bishop G, Hamberger B, Höfer R, Paquette S, Werck-Reichhart D. 2011. Cytochromes P450. *The Arabidopsis Book*: e0144.

Barnes HJ, Arlotto MP, Waterman MR. 1991. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in *Escherichia coli*. *Proc Natl Acad Sci U S A,* **88**: 5597-601.

Barrett RD, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol Evol,* **23**: 38-44.

Barrett SCH. 2008. Major evolutionary transitions in flowering plant reproduction: An overview. *International Journal Of Plant Sciences,* **169**: 1-5.

Bassett MJ, Miklas PN. 2007. A new gene, *bic*, with pleiotropic effects (with T P V) for bicolor flowers and dark olive brown seed coat in common bean. *Journal of the American Society for Horticultural Science,* **132**: 352-356.

Bässler M. 1973. Revision der eurasiatischen Arten von *Lathyrus* L. Sect. *Orobus* (L.) Gren. et Godr. *Feddes Repertorium,* **84**: 329-447.

Bateson W. 1905. *Notes on the Preogress of Mendelian Studies*.

Bateson W. 1913. Mendel's principles of heredity. Chapter V, p88. Cambridge: University Press.

Bateson W, Saunders ER, Punnett RC. 1906. Further experiments on inheritance in sweet peas and stocks: Preliminary account. *Proceedings of the Royal Society of London Series B,* **77**: 236-238.

Bateson W, Saunders ER, Punnett RC, Hurst CC, Wheldale M, Marryat DCE, Sollas IBJ. 1902. *Experimental studies in the physiology of heredity.* London: Harrison.

Bateson W, Waunders ER, Punnett RC. 1909. Experimental studies in the physiology of heredity. *Zeitschrift für Induktive Abstammungs- und Vererbungslehre,* **2**: 17-19.

Beale GH, Robinson GM, Robinson R, Scott-Moncrieff R. 1939. Genetics and chemistry of flower

colour variation in *Lathyrus odoratus*. *Journal of Genetics,* **37**: 375.

Beatty J. 2010. Reconsidering the Importance of Chance Variation. In: Pigliucci M, Müller GB, eds. *Evolution, the Extended Synthesis*: The MIT Press.

Benveniste I, Lesot A, Hasenfratz MP, Kochs G, Durst F. 1991. Multiple forms of NADPH-cytochrome P450 reductase in higher plants. *Biochemical and Biophysical Research Communications,* **177**: 105-112.

Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C. 2000. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell,* **12**: 2383-2394.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature,* **426**: 176-178.

Brahim NB, Combes D, Marrakchi M. 2001. Autogamy and allogamy in genus *Lathyrus*. *Lathyrus Lathyrism Newsletter,* **2**: 21-26.

Brahim NB, Salhi A, Chtourou N, Combes D, Marrakchi M. 2002. Isozymic polymorphism and phylogeny of 10 *Lathyrus* species. *Genetic Resources and Crop Evolution,* **49**: 429-438.

Briscoe AD, Chittka L. 2001. The evolution of color vision in insects. *Annual Review of Entomology,* **46**: 471-510.

Broich SL. 2007. New combinations in north american *Lathyrus* and *Vicia* (Fabaceae: Faboideae: Fabeae). *Madroño,* **54**: 63-71.

Broich SL. 2009. *Lathyrus,* manuscript prepared for Flora of North America.

Brugliera F, Tull D, Holton TA, Karan M, Treloar N, Simpson K, Skurczynska J, Mason JG. 2000. Introduction of a cytochrome b5 enhances the activity of flavonoid 3',5' hydroxylase (a cytochrome P450) in transgenic carnation. *Sixth International Congress of Plant Molecular Biology.* University of Laval, Quebec.

Caldwell MM, Bornman JF, Ballare CL, Flint SD, Kulandaivelu G. 2007. Terrestrial ecosystems, increased solar ultraviolet radiation, and interactions with bother climate change factors. *Photochemical & Photobiological Sciences,* **6**: 252-266.

Carey CC, Strahle JT, Selinger DA, Chandler VL. 2004. Mutations in the *pale aleurone color1* regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar *TRANSPARENT TESTA GLABRA1* gene in *Arabidopsis thaliana*. *Plant Cell,* **16**: 450-64.

Cavaiuolo M, Cocetta G, Ferrante A. 2013. The antioxidants changes in ornamental flowers during development and senescence. *Antioxidants,* **2**: 132.

Chittka L, Raine NE. 2006. Recognition of flowers by pollinators. *Current Opinion in Plant Biology,* **9**: 428-435.

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

Chrtkova-Zertova A, van der Maesen LJG, Rechinger KH. 1979. *Flora Iranica. Papillinoid I.* Austria: Akademische Druck und verlagsantalt Graz.

Clegg MT, Durbin ML. 2003. Tracing floral adaptations from ecology to molecules. *Nat Rev Genet,* **4**: 206-215.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J,* **16**: 735-43.

Comba L, Corbet SA, Hunt H, Outram S, Parker JS, Glover BJ. 2000. The role of genes

influencing the corolla in pollination of *Antirrhinum majus*. *Plant, Cell & Environment,* **23**: 639-647.

Cone KC, Cocciolone SM, Burr FA, Burr B. 1993. Maize anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant. *Plant Cell,* **5**: 1795-1805.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci,* **279**: 5039-47.

Cronk QC. 2009. Evolution in reverse gear: the molecular basis of loss and reversal. *Cold Spring Harb Symp Quant Biol,* **74**: 259-66.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods,* **9**: 772.

Davies K, Schwinn K, Deroles S, Manson D, Lewis D, Bloor S, Bradley JM. 2003. Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. *Euphytica,* **131**: 259-268.

de Vetten N, Quattrocchio F, Mol J, Koes R. 1997. The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev,* **11**: 1422-34.

de Vetten N, ter Horst J, van Schaik HP, de Boer A, Mol J, Koes R. 1999. A cytochrome b(5) is required for full activity of flavonoid 3',5'-hydroxylase, a cytochrome P450 involved in the formation of blue flower colors. *Proceedings of the National Academy of Sciences of the United States of America,* **96**: 778-783.

Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution,* **24**: 332-340.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. *Evolution,* **64**: 2044-2054.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15

Drummond A, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology,* **7**: 214.

Durbin ML, Lundy KE, Morrell PL, Torres-Martinez CL, Clegg MT. 2003. Genes that determine flower color: the role of regulatory changes in the evolution of phenotypic adaptations. *Molecular Phylogenetics and Evolution,* **29**: 507-518.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res,* **32**: 1792-7.

Ellis AG, Johnson SD. 2009. The evolution of floral variation without pollinator shifts in *Gorteria diffusa* (Asteraceae). *American Journal of Botany,* **96**: 793-801.

Falginella L, Castellarin S, Testolin R, Gambetta G, Morgante M, Di Gaspero G. 2010. Expansion and subfunctionalisation of flavonoid 3',5'-hydroxylases in the grapevine lineage. *BMC Genomics,* **11**: 562.

Farzad M, Griesbach R, Hammond J, Weiss MR, Elmendorf HG. 2003. Differential expression of three key anthocyanin biosynthetic genes in a color-changing flower, *Viola cornuta* cv. Yesterday, Today and Tomorrow. *Plant Science,* **165**: 1333-1342.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology Evolution and Systematics,* **35**: 375-403.

Ferrer JL, Austin MB, Stewart C, Jr., Noel JP. 2008. Structure and function of enzymes involved

in the biosynthesis of phenylpropanoids. *Plant Physiol Biochem,* **46**: 356-70.

Fineblum WL, Rausher MD. 1997. Do floral pigmentation genes also influence resistance to enemies? The *W* locus in *Ipomoea purpurea*. *Ecology,* **78**: 1646-1654.

Fisher RA, Ford EB. 1950. The Sewall Wright effect. *Heredity,* **4**: 117-119.

Fleming WM. 1925. *Inheritance in sweet peas,* Master's thesis, University of British Columbia.

Forkmann G. 1991. Flavonoids as flower pigments: The formation of the natural spectrum and its extension by genetic engineering. *Plant Breeding,* **106**: 1-26.

Forkmann G, Ruhnau B. 1987. Distinct substrate-specificity of dihydroflavonol 4-reductase from flowers of *Petunia hybrida*. *Zeitschrift Fur Naturforschung C-a Journal Of Biosciences,* **42**: 1146-1148.

Fukui Y, Tanaka Y, Kusumi T, Iwashita T, Nomoto K. 2003. A rationale for the shift in colour towards blue in transgenic carnation flowers expressing the flavonoid 3',5'-hydroxylase gene. *Phytochemistry,* **63**: 15-23.

Fulton M, Hodges SA. 1999. Floral isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proceedings of the Royal Society B: Biological Sciences,* **266**: 2247-2247.

Gerats T, Strommer J. 2009. *Petunia* evolutionary, developmental and physiological genetics. New York, NY: Springer New York.

Glover BJ, Martin C. 1998. The role of petal cell shape and pigmentation in pollination success in *Antirrhinum majus*. *Heredity,* **80**: 778-784.

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. 2008. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant Journal,* **53**: 814-827.

Goodrich J, Carpenter R, Coen ES. 1992. A common gene regulates pigmentation pattern in diverse plant species. *Cell,* **68**: 955-964.

Gou JY, Felippes FF, Liu CJ, Weigel D, Wang JW. 2011. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell,* **23**: 1512-22.

Gould KS. 2004. Nature's Swiss army knife: The diverse protective roles of anthocyanins in leaves. *Journal of Biomedicine and Biotechnology*: 314-320.

Griffiths A, Ganders F. 1983. *Wildflower genetics: a field guide for British Columbia and the Pacific Northwest*: Flight Press.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol,* **52**: 696-704.

Guo J, Han W, Wang MH. 2008. Ultraviolet and environmental stresses involved in the induction and regulation of anthocyanin biosynthesis: A review. *African Journal of Biotechnology,* **7**: 4966-4972.

Gutierrez JF, Vaquero F, Vences FJ. 1994. Allopolyploid vs. autopolyploid origins in the genus *Lathyrus* (Leguminosae). *Heredity,* **73**: 29-40.

Hammett KRW, Murray BG, Markham KR, Hallett IC. 1994. Interspecific hybridization between *Lathyrus odoratus* and *L. belinensis*. *International Journal of Plant Sciences,* **155**: 763-771.

Harborne JB. 1988. *The Flavonoids: Advances in Research since 1980*: Springer US.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics,* **24**: 129-131.

Hatcher PE, Paul ND. 1994. The effect of elevated uv-b radiation on herbivory of pea by *Autographa gamma*. *Entomologia Experimentalis Et Applicata,* **71**: 227-233.

He H, Ke H, Keting H, Qiaoyan X, Silan D. 2013. Flower colour modification of chrysanthemum by suppression of *F3'H* and overexpression of the exogenous *Senecio cruentus F3'5'H* gene. *PLoS ONE,* **8**: e74395.

Heath TA, Hedtke SM, Hillis DM. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal Of Systematics And Evolution,* **46**: 239-257.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution,* **27**: 570-580.

Hellens RP, Moreau C, Lin-Wang K, Schwinn KE, Thomson SJ, Fiers MWEJ, Frew TJ, Murray SR, Hofer JMI, Jacobs JME, Davies KM, Allan AC, Bendahmane A, Coyne CJ, Timmerman-Vaughan GM, Ellis THN. 2010. Identification of Mendel's white flower character. *PLoS ONE,* **5**: e13230.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007a. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell,* **19**: 779-790.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007b. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell,* **19**: 779-90.

Hodges SA, Derieg NJ. 2009. Adaptive radiations: From field to genomic studies. *Proceedings of the National Academy of Sciences of the United States of America,* **106**: 9947-9954.

Hoekema A, Kastelein RA, Vasser M, de Boer HA. 1987. Codon replacement in the *PGK1* gene of *Saccharomyces cerevisiae*: experimental approach to study the role of biased codon usage in gene expression. *Mol Cell Biol,* **7**: 2914-24.

Hofer J, Turner L, Moreau C, Ambrose M, Isaac P, Butcher S, Weller J, Dupin A, Dalmais M, Le Signor C, Bendahmane A, Ellis N. 2009. *Tendril-less* regulates tendril formation in pea leaves. *The Plant Cell,* **21**: 420-428.

Holton TA, Cornish EC. 1995. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell,* **7**: 1071-1083.

Hopkins R, Rausher MD. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature,* **469**: 411-414.

Hoshino A, Morita Y, Choi JD, Saito N, Toki K, Tanaka Y, Iida S. 2003. Spontaneous mutations of the *flavonoid 3'-hydroxylase* gene conferring reddish flowers in the three morning glory species. *Plant Cell Physiol,* **44**: 990-1001.

Inagaki Y, Johzuka-Hisatomi Y, Mori T, Takahashi S, Hayakawa Y, Peyachoknagul S, Ozeki Y, Iida S. 1999. Genomic organization of the genes encoding dihydroflavonol 4-reductase for flower pigmentation in the Japanese and common morning glories. *Gene,* **226**: 181-188.

Jenkins GI, Fuglevand G, Christle JM. 1997. UV-B perception and signal transduction. In: Lumsden PJ, ed. *Plants and UV-B : responses to environmental change*. Cambridge: Cambridge Univ. Press.

Johnson ET, Ryu S, Yi H, Shin B, Cheong H, Choi G. 2001. Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. *The Plant Journal,* **25**: 325-333.

Johnson ET, Yi H, Shin B, Oh B-J, Cheong H, Choi G. 1999. *Cymbidium hybrida* dihydroflavonol 4-reductase does not efficiently reduce dihydrokaempferol to produce orange pelargonidin-type anthocyanins. *The Plant Journal,* **19**: 81-85.

Johnson SD. 2006. Pollinator-driven speciation in plants. In: Harder LD, Barrett SCH, eds. *The ecology and evolution of flowers*. Oxford, UK: Oxford University Press.

Kalisz S, Vogler DW, Hanley KM. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. *Nature,* **430**: 884-887.

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekura-Sakakibara K, Togami J, Pigeaire A, Tao GQ, Nehra NS, Lu CY, Dyson BK, Tsuda S, Ashikari T, Kusumi T, Mason JG, Tanaka Y. 2007. Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. *Plant and Cell Physiology,* **48**: 1589-1600.

Kearney JP. 1993. *Wild Lathyrus species as genetic resource for improvement of Grasspea (L.sativus),* University of Southampton.

Kenicer GJ, Kajita T, Pennington RT, Murata J. 2005. Systematics and biogeography of *Lathyrus* (Leguminosae) based on internal transcribed spacer and cpDNA sequence data. *American Journal of Botany,* **92**: 1199-1209.

Kevan PG, Baker HG. 1983. Insects as flower visitors and pollinators. *Annual Review of Entomology,* **28**: 407-453.

Kislev ME. 1989. Origins of the cultivation of *Lathyrus sativus* and *L. cicera* (Fabaceae). *Economic Botany,* **43**: 262-270.

Knuth P, Müller H, Davis JRA. 1906. *Handbook of flower pollination: based upon Hermann Müller's work 'The fertilisation of flowers by insects'*. Oxford: Clarendon Press.

Koes RE, Quattrocchio F, Mol JNM. 1994. The flavonoid biosynthetic pathway in plants: Function and evolution. *BioEssays,* **16**: 123-132.

Kovinich N, Kayanja G, Chanoca A, Riedl K, Otegui M, Grotewold E. 2014. Not all anthocyanins are born equal: distinct patterns induced by stress in *Arabidopsis*. *Planta,* **240**: 931-940.

Kupicha FK. 1983. The infrageneric structure of lathyrus. *Notes from the Royal Botanic Garden Edinburgh,* **41**: 209-244.

Landry LG, Chapple CCS, Last RL. 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-b injury and oxidative damage. *Plant Physiology,* **109**: 1159-1166.

Lavin M, Herendeen PS, Wojciechowski MF. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Syst Biol,* **54**: 575-94.

Lawrence WJC, Price JR, Robinson GM, Robinson R. 1941. The distribution of anthocyanins in flowers, fruits and leaves. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences,* **230**: 149-178.

Lewis G, Schrire B, Mackinder B, Lock M. 2005. *Legumes of the world*: Royal Botanic Gardens, Kew.

Lorenc-Kukula K, Amarowicz R, Oszmianski J, Doermann P, Starzycki M, Skala J, Zuk M, Kulma A, Szopa J. 2005. Pleiotropic effect of phenolic compounds content increases in transgenic flax plant. *J Agric Food Chem,* **53**: 3685-92.

Loreti E, Povero G, Novi G, Solfanelli C, Alpi A, Perata P. 2008. Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in

Arabidopsis. *New Phytologist*, **179**: 1004-1016.

Ludwig SR, Wessler SR. 1990. Maize R gene family: tissue-specific helix-loop-helix proteins. *Cell*, **62**: 849-51.

Maddison WP, Maddison DR. 2011. Mesquite: a modular system for evolutionary analysis.

Markham KR. 1982. Techniques of flavonoid identification. London: Academic Press.

Martin C, Gerats T. 1993. Control of pigment biosynthesis genes during petal development. *The Plant Cell*, **5**: 1253-1264.

Martins TR, Berg JJ, Blinka S, Rausher MD, Baum DA. 2013. Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in *Clarkia gracilis* (Onagraceae). *New Phytol*, **197**: 958-69.

Mayr E, Provine WB. 1980. *The evolutionary synthesis: Perspectives on the unification of biology.* Cambridge, MA: Harvard UP.

Mazer SJ, Paz H, Bell MD. 2004. Life history, floral development and mating system in *Clarkia xantiana* (Onagraceae): Do floral and whole-plant rates of development evolve independently. *American Journal of Botany*, **91**: 2041-2050.

McEwen JR, Vamosi JC. 2010. Floral colour versus phylogeny in structuring subalpine flowering communities. *Proc Biol Sci*, **277**: 2957-65.

Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T. 2008. Flowering-time genes modulate meristem determinacy and growth form in *Arabidopsis thaliana*. *Nat Genet*, **40**: 1489-92.

Menting JG, Cornish E, Scopes RK. 1994a. Purification and partial characterization of NADPH-cytochrome c reductase from *Petunia hybrida* flowers. *Plant Physiol*, **106**: 643-50.

Menting JGT, Scopes RK, Stevenson TW. 1994b. Characterization of flavonoid 3',5'-hydroxylase in microsomal membrane fraction of *Petunia hybrida* flowers. *Plant Physiology*, **106**: 633-642.

Meyer P, Heidmann I, Forkmann G, Saedler H. 1987. A new petunia flower color generated by transformation of a mutant with a maize gene. *Nature*, **330**: 677-678.

Mizutani M, Ohta D. 1998. Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*. Gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol*, **116**: 357-67.

Mol J, Grotewold E, Koes R. 1998. How genes paint flowers and seeds. *Trends in Plant Science*, **3**: 212-217.

Moller IM, Lin W. 1986. Membrane-bound NAD(P)H dehydrogenases in higher plant cells. *Annual Review of Plant Physiology*, **37**: 309-334.

Morales LO, Tegelberg R, Brosché M, Keinänen M, Lindfors A, Aphalo PJ. 2010. Effects of solar UV-A and UV-B radiation on gene expression and phenolic accumulation in *Betula pendula* leaves. *Tree Physiology*, **30**: 923-934.

Moreau C, Ambrose MJ, Turner L, Hill L, Ellis THN, Hofer JMI. 2012. The *b* gene of pea (*Pisum sativum* L.) encodes a defective flavonoid 3' 5' hydroxylase and confers pink flower colour. *Plant Physiology*.

Mori M, Yoshida K, Ishigaki Y, Matsunaga T, Nikaido O, Kameda K, Kondo T. 2005. UV-B protective effect of a polyacylated anthocyanin, HBA, in flower petals of the blue morning glory, *Ipomoea tricolor* cv. Heavenly Blue. *Bioorg Med Chem*, **13**: 2015-20.

Muchhala N, Johnsen S, Smith SD. 2014. Competition for hummingbird pollination shapes flower

color variation in andean Solanaceae. *Evolution*, **68**: 2275-2286.

Murray BG, Hammett KRW. 1998. New sweet pea (*Lathyrus odoratus*) cultivars via interspecific hybridization with wild *Lathyrus* species. *Third International Symposium on New Floricultural Crops - Proceedings*: 115-118.

Nabhan AR, Sarkar IN. 2011. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. *Briefings in Bioinformatics*.

Nakatsuka T, Nishihara M, Mishiba K, Hirano H, Yamamura S. 2006. Two different transposable elements inserted in *flavonoid 3',5'-hydroxylase* gene contribute to pink flower coloration in *Gentiana scabra*. *Molecular Genetics and Genomics*, **275**: 231-241.

Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L. 2000. The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* siliques. *Plant Cell*, **12**: 1863-78.

Ojeda DI, Santos-Guerra A, Oliva-Tejera F, Valido A, Xue X, Marrero A, Caujapé-Castells J, Cronk Q. 2013. Bird-pollinated Macaronesian *Lotus* (Leguminosae) evolved within a group of entomophilous ancestors with post-anthesis flower color change. *Perspectives in Plant Ecology, Evolution and Systematics*, **15**: 193-204.

Okinaka Y, Shimada Y, Nakano-Shimada R, Ohbayashi M, Kiyokawa S, Kikuchi Y. 2003. Selective accumulation of delphinidin derivatives in tobacco using a putative *flavonoid 3',5'-hydroxylase* cDNA from *Campanula medium*. *Biosci Biotechnol Biochem*, **67**: 161-5.

Onslow MW. 2014. *The anthocyanin pigments of plants*: Cambridge University Press.

Ostevik KL, Moyers BT, Owens GL, Rieseberg LH. 2012. Parallel ecological speciation in plants? *International Journal of Ecology*, **2012**: 17.

Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature*, **401**: 877-884.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**: 289-90.

Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H. 1987. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J*, **6**: 3553-8.

Pecket RC. 1960. The nature of the variation in flower colour in the genus *Lathyrus*. *New Phytologist*, **59**: 138-144.

Pérez F, Arroyo MTK, Medel R, Hershkovitz MA. 2006. Ancestral reconstruction of flower morphology and pollination systems in Schizanthus (Solanaceae). *American Journal of Botany*, **93**: 1029-1038.

Perret M, Chautems A, Spichiger R, Kite G, Savolainen V. 2003. Systematics and evolution of tribe *Sinningieae* (Gesneriaceae): evidence from phylogenetic analyses of six plastid DNA regions and nuclear ncpGS. *American Journal of Botany*, **90**: 445-460.

Pinheiro J, Bates D, DebRoy S, Sarkar D. 2013. R Development Core Team (2012) nlme: linear and nonlinear mixed effects models. R package version 3.1-103. *R Foundation for Statistical Computing, Vienna*.

Plotkin JB, Kudla G. 2011. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*, **12**: 32-42.

Punnett RC. 1907. Mendelism. 2nd ed ed. Cambridge: Bowes and Bowes.

Punnett RC. 1923. Linkage in the sweet pea (*Lathyrus odoratus*). *Journal of Genetics*, **13**: 101-U9.

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

Quattrocchio F, Wing JF, Leppen HTC, Mol JNM, Koes RE. 1993. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *The Plant Cell,* **5**: 1497-1512.

Quattrocchio F, Wing JF, Va K, De, Woude, Mol JNM, Koes R. 1998. Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. *The Plant Journal,* **13**: 475-488.

Ramsay NA, Glover BJ. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci,* **10**: 63-70.

Rat'kin AV, Tarasov VA. 2010. Genetic control of biosynthesis of anthocyans in sweet pea (*Lathyrus odoratus* L.) flowers. *Russian Journal of Genetics,* **46**: 431-438.

Rausher MD. 2006. The evolution of flavonoids and their genes. In: Grotewold E, ed. *The Science of Flavonoids*: Springer New York.

Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal Of Plant Sciences,* **169**: 7-21.

Renoult J, Kjellberg F, Grout C, Santoni S, Khadari B. 2009. Cyto-nuclear discordance in the phylogeny of Ficus section Galoglychia and host shifts in plant-pollinator associations. *BMC Evolutionary Biology,* **9**: 248.

Ritland K, Newton C, Marshall HD. 2001. Inheritance and population structure of the white-phased "Kermode" black bear. *Current Biology,* **11**: 1468-1472.

Ross JJ, Murfet IC. 1988. Flowering and branching in *Lathyrus odoratus* L.: loci *sp* and *b*. *Theoretical and Applied Genetics,* **76**: 941-946.

Rupasinghe S, Baudry J, Schuler MA. 2003. Common active site architecture and binding strategy of four phenylpropanoid P450s from *Arabidopsis thaliana* as revealed by molecular modeling. *Protein Engineering,* **16**: 721-731.

Rupasinghe S, Schuler MA. 2006. Homology modeling of plant cytochrome P450s. *Phytochemistry Reviews,* **5**: 473-505.

Sahin A, Genc H, Bagci E. 2000. Cytotaxonomic investigations on some *Lathyrus* L. species growing in eastern Mediterranean and southern Aegean regions II. *Acta Botanica Gallica,* **147**: 243-256.

Sato M, Kawabe T, Hosokawa M, Tatsuzawa F, Doi M. 2011. Tissue culture-induced flower-color changes in *Saintpaulia* caused by excision of the transposon inserted in the *flavonoid 3', 5' hydroxylase* (*F3'5'H*) promoter. *Plant Cell Reports,* **30**: 929-939.

Schaefer H, Hechenleitner P, Santos-Guerra A, de Sequeira MM, Pennington RT, Kenicer G, Carine M. 2012. Systematics, biogeography, and character evolution of the legume tribe Fabeae with special focus on the middle-Atlantic island lineages. *BMC Evolutionary Biology,* **12**: 250.

Schoenbohm C, Martens S, Eder C, Forkmann G, Weisshaar B. 2000. Identification of the *Arabidopsis thaliana flavonoid 3'-hydroxylase* gene and functional expression of the encoded P450 enzyme. *Biological Chemistry,* **381**: 749-753.

Schwinn K, Davies K, Alm V, Mackay S, Martin C. 2001. Regulation of anthocyanin biosynthesis

in *Antirrhinum*. In: Sorvari S, Karhu S, Kanervo E, Pihakaski S, eds. *Proceedings of the 4th International Symposium on in Vitro Culture and Horticultural Breeding*. Leuven 1: International Society Horticultural Science.

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell,* **18**: 831-51.

Searles PS, Flint SD, Caldwell MM. 2001. A meta analysis of plant field studies simulating stratospheric ozone depletion. *Oecologia,* **127**: 1-10.

Seitz C, Ameres S, Forkmann G. 2007. Identification of the molecular basis for the functional difference between flavonoid 3′-hydroxylase and flavonoid 3′,5′-hydroxylase. *FEBS Letters,* **581**: 3429-3434.

Seitz C, Eder C, Deiml B, Kellner S, Martens S, Forkmann G. 2006. Cloning, functional identification and sequence analysis of *flavonoid 3'-hydroxylase* and *flavonoid 3',5'-hydroxylase* cDNAs reveals independent evolution of *flavonoid 3',5'-hydroxylase* in the Asteraceae family. *Plant Molecular Biology,* **61**: 365-381.

Shang Y, Schwinn K, Bennett M, Hunter D, Waugh T, Pathirana N, Brummell D, Jameson P, Davies K. 2007. Methods for transient assay of gene function in floral tissues. *Plant Methods,* **3**: 1.

Shehadeh A. 2011. *Ecogeographic, genetic and taxonomic studies of the genus Lathyrus L.,* PhD Thesis, University of Birmingham.

Shi MZ, Xie DY. 2010. Features of anthocyanin biosynthesis in *pap1-D* and wild-type *Arabidopsis thaliana* plants grown in different light intensity and culture media conditions. *Planta,* **231**: 1385-400.

Smith SD, Rausher MD. 2011. Gene loss and parallel evolution contribute to species difference in flower color. *Molecular Biology and Evolution*.

Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a flower color transition. *Molecular biology and evolution,* **30**: 602-12.

Sobel JM, Streisfeld MA. 2013. Flower color as a model system for studies of plant evo-devo. *Front Plant Sci,* **4**: 321.

Soltis DE, Kuzoff RK. 1995. Discordance between nuclear and chloroplast phylogenies in the Heuchera group (Saxifragaceae). *Evolution,* **49**: 727-742.

Spelt C, Quattrocchio F, Mol J, Koes R. 2002. *Anthocyanin1* of *Petunia* controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. *Plant Cell,* **14**: 2121-2135.

Spelt C, Quattrocchio F, Mol JNM, Koes R. 2000. *Anthocyanin1* of *Petunia* encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *The Plant Cell Online,* **12**: 1619-1631.

Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Genet,* **14**: 751-764.

Stich K, Eidenberger T, Wurst F, Forkmann G. 1992. Enzymatic conversion of dihydroflavonols to flavan-3,4-diols using flower extracts of *Dianthus caryophyllus* L. (carnation). *Planta,* **187**: 103-108.

Stracke R, Ishihara H, Barsch GHA, Mehrtens F, Niehaus K, Weisshaar B. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation

in different parts of the *Arabidopsis thaliana* seedling. *Plant Journal,* **50**: 660-677.

Strauss SY, Whittall JB. 2006. Non-pollinator agents of selection on floral traits. In: Harder LB, Barrett SCH, eds. *Ecology and Evolution of Flowers*. Oxford: Oxford University Press.

Streisfeld MA, Kohn JR. 2005. Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution,* **59**: 2548-59.

Streisfeld MA, Kohn JR. 2007. Environment and pollinator-mediated selection on parapatric floral races of *Mimulus aurantiacus*. *J Evol Biol,* **20**: 122-32.

Streisfeld MA, Liu D, Rausher MD. 2011. Predictable patterns of constraint among anthocyanin-regulating transcription factors in *Ipomoea*. *New Phytol,* **191**: 264-74.

Streisfeld MA, Rausher MD. 2009. Altered trans-regulatory control of gene expression in multiple anthocyanin genes contributes to adaptive flower color evolution in *Mimulus aurantiacus*. *Molecular Biology and Evolution,* **26**: 433-444.

Takahashi R, Dubouzet J, Matsumura H, Yasuda K, Iwashina T. 2010. A new allele of flower color gene *W1* encoding *flavonoid 3'5'-hydroxylase* is responsible for light purple flowers in wild soybean *Glycine soja*. *BMC Plant Biology,* **10**: 155.

Takeda K, Osakabe A, Saito S, Furuyama D, Tomita A, Kojima Y, Yamadera M, Sakuta M. 2005. Components of protocyanin, a blue pigment from the blue flowers of *Centaurea cyanus*. *Phytochemistry,* **66**: 1607-1613.

Tanaka Y. 2006. Flower colour and cytochromes P450. *Phytochemistry Reviews,* **5**: 283-291.

Tanaka Y, Brugliera F, Chandler S. 2009. Recent progress of flower colour modification by biotechnology. *International Journal of Molecular Sciences,* **10**: 5350-5369.

Tanaka Y, Fukui Y, Fukuchimizutani M, Holton TA, Higgins E, Kusumi T. 1995. Molecular cloning and characterization of *Rosa hybrida dihydroflavonol 4-reductase* gene. *Plant And Cell Physiology,* **36**: 1023-1031.

Tanaka Y, Ohmiya A. 2008. Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. *Curr Opin Biotechnol,* **19**: 190-7.

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K. 2005. Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. *Plant J,* **42**: 218-35.

Uimari A, Strommer J. 1997. Myb26: a MYB-like protein of pea flowers with affinity for promoters of phenylpropanoid genes. *Plant Journal,* **12**: 1273-1284.

Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D. 1997. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J Biol Chem,* **272**: 19176-86.

Vain P, Keen N, Murillo J, Rathus C, Nemes C, Finer J. 1993. Development of the particle inflow gun. *Plant Cell, Tissue and Organ Culture,* **33**: 237-246.

Varadarajan J, Guilleminot J, Saint-Jore-Dupas C, Piegu B, Chaboute ME, Gomord V, Coolbaugh RC, Devic M, Delorme V. 2010. *ATR3* encodes a diflavin reductase essential for *Arabidopsis* embryo development. *New Phytol,* **187**: 67-82.

Wake DB. 1991. Homoplasy - the result of natural selection, or evidence of design limitations. *American Naturalist,* **138**: 543-567.

Waser NM. 1983. The adaptive nature of floral traits: ideas and evidence. In: Real LA, ed. *Pollination*

Biology: Academic Press.

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in Pollination Systems, and Why it Matters. *Ecology*, **77**: 1043-1060.

Waser NM, Price MV. 1983. Pollinator behaviour and natural selection for flower colour in *Delphinium nelsonii*. *Nature,* **302**: 422-424.

Wei YZ, Hu FC, Hu GB, Li XJ, Huang XM, Wang HC. 2011. Differential expression of anthocyanin biosynthetic genes in relation to anthocyanin accumulation in the pericarp of *Litchi chinensis* Sonn. *PLoS One,* **6**: e19455.

Weiss MR. 1995. Floral color change: A widespread functional convergence. *American Journal of Botany,* **82**: 167-185.

Welsch R, Arango J, Bär C, Salazar B, Al-Babili S, Beltrán J, Chavarriaga P, Ceballos H, Tohme J, Beyer P. 2010. Provitamin a accumulation in cassava (*manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *The Plant Cell,* **22**: 3348-3356.

Wessinger CA, Rausher MD. 2012. Lessons from flower colour evolution on targets of selection. *Journal of Experimental Botany,* **63**: 5741-5749.

Wessinger CA, Rausher MD. 2014. Predictability and irreversibility of genetic changes associated with flower color evolution in *Penstemon barbatus*. *Evolution,* **68**: 1058-70.

Wessinger CA, Rausher MD. 2015. Ecological transition predictably associated with gene degeneration. *Molecular Biology and Evolution,* **32**: 347-354.

Westerkamp C. 1993. The cooperation between the asymmetric flower of *Lathyrus latifolius* (Fabaceae-Vicieae) and its visitors. *Phyton-Annales Rei Botanicae,* **33**: 121-137.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. *Molecular Ecology,* **15**: 4645-4657.

Wiering H. 1974. Genetics of flower color in *Petunia hybrida Genen en Phaenen*, **17**: 117-134.

Wilson P, Castellanos MC, Hogue JN, Thomson JD, Armbruster WS. 2004. A multivariate search for pollination syndromes among penstemons. *Oikos,* **104**: 345-361.

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

Woodall GS, Stewart GR. 1998. Do anthocyanins play a role in UV protection of the red juvenile leaves of *Syzygium*? *Journal of Experimental Botany*, **49**: 1447-1450.

Woollacott CM. 2010. *The genetic control of petal morphology in the sweet pea (Lathyrus odoratus),* Master's thesis, University of British Columbia, Vancouver.

Wu CA, Streisfeld MA, Nutter LI, Cross KA. 2013. The genetic basis of a rare flower color polymorphism in *Mimulus lewisii* provides insight into the repeatability of evolution. *PLoS ONE,* **8**: e81173.

Wylie CD, Darlington AP. 1955. *Chromosome atlas of flowering plants.*: George Allen & Unwin.

Yan F, Di S, Rojas Rodas F, Rodriguez Torrico T, Murai Y, Iwashina T, Anai T, Takahashi R. 2014. Allelic variation of soybean flower color gene *W4* encoding dihydroflavonol 4-reductase 2. *BMC Plant Biology,* **14**: 58.

Yuan YW, Sagawa JM, Young RC, Christensen BJ, Bradshaw HD, Jr. 2013. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics,* **194**: 255-63.

Zabala G, Vodkin LO. 2007. A rearrangement resulting in small tandem repeats in the *F3'5'H* gene of white flower genotypes is associated with the soybean *W1* locus. *Crop Science,* **47**: S113-S124.

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A. 2003. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development,* **130**: 4859-69.

Zufall RA, Rausher MD. 2003. The genetic basis of a flower color polymorphism in the common morning glory (*Ipomoea purpurea*). *J Hered,* **94**: 442-8.

Zwickl DJ, Hillis DM. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology,* **51**: 588-598.