The role of pollinator-mediated selection in the maintenance of a flower colour polymorphism in an *Antirrhinum majus* hybrid zone

by

Thomas James Ellis

18th February 2016

A thesis presented to the faculty of the graduate school of the Institute of Science and Technology Austria, Klosterneuburg, Austria in partial fulfilment of the requirements for the degree Doctor of Philosophy.

[redacted]

Supervisor: Nick Barton, IST Austria, Klosterneuburg

[redacted]

Committee Member: Sylvia Crémer, IST Austria, Klosterneuburg

[redacted]

Committee Member: Magnus Nordborg, Gregor Mendel [redacted] itute, Vienna

Programme Chair: Gašper Tkačik, IST Austria, Klosterneuburg

[redacted]

I hereby declare that this dissertation is my own work, and it does not contain other peoples work without this being so stated; and this thesis does not contain my previous work without this being stated, and that the bibliography contains all the literature that I used in writing the dissertation, and that all references refer to this bibliography. I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Thomas James Ellis, 19th February 2016

Acknowledgements

I am indebted to many people for their support during my PhD, but I particularly wish to thank Nick Barton for his guidance and intuition, and for encouraging me to take the time to look beyond the immediate topic of my PhD to understand the broader context. I am also especially grateful to David Field his bottomless patience, invaluable advice on experimental design, analysis and scientific writing, and for tireless work on the population surveys and genomic work without most of my thesis could not have happened.

It has been a pleasure to work with the combined strengths of the groups at The John Innes Centre, University of Toulouse and IST Austria. Thanks to Enrico Coen and his group for hosting me in Norwich in 2011 and especially for setting up the tag experiment.

I thank David Field, Desmond Bradley and Maria Clara Melo-Hurtado for organising field collections, as well as Monique Burrus and Christophe Andalo and a large number of volunteers for their efforts helping with the field work. Furthermore I thank Coline Jaworski for providing seeds and for her input into the design of the experimental arrays, and Matthew Couchman for maintaining the database of.

In addition to those mentioned above, I am grateful to Melinda Pickup, Spencer Barrett, and four anonymous reviewers for their insightful comments on sections of this manuscript. I also thank Jana Porsche for her efforts in tracking down the more obscure references for chapter 5, and Jon Bollback for his advice about the analysis.

I am indebted to Jon Ågren for his patience whilst I finished this thesis, and to Sylvia Cremer and Magnus Nordborg for taking the time to read and evaluate the thesis given a shorter deadline than was fair.

A very positive aspect of my PhD has been the supportive atmosphere of IST. In particular, I have come to appreciate the enormous support from

our group assistants Nicole Hotzy, Julia Asimakis, Christine Ostermann and Jerneja Beslagic. I also thank Christian Chaloupka and Stefan Hipfinger for their enthusiasm and readiness to help where possible in setting up our greenhouse and experiments.

Finally I thank Barbie for her love and her patience whilst this thesis swallowed my life.

Abstract

Hybrid zones represent evolutionary laboratories, where recombination brings together alleles in combinations which have not previously been tested by selection. This provides an excellent opportunity to test the effect of molecular variation on fitness, and how this variation is able to spread through populations in a natural context. The snapdragon *Antirrhinum majus* is polymorphic in the wild for two loci controlling the distribution of yellow and magenta floral pigments. Where the yellow *A. m. striatum* and the magenta *A. m. pseudomajus* meet along a valley in the Spanish Pyrenees they form a stable hybrid zone. Alleles at these loci recombine to give striking transgressive variation for flower colour. The sharp transition in phenotype over ∼1km implies strong selection maintaining the hybrid zone. An indirect assay of pollinator visitation in the field found that pollinators forage in a positive-frequency dependent manner on *Antirrhinum,* matching previous data on fruit set. Experimental arrays and paternity analysis of wild-pollinated seeds demonstrated assortative mating for pigmentation alleles, and that pollinator behaviour alone is sufficient to explain this pattern. Selection by pollinators should be sufficiently strong to maintain the hybrid zone, although other mechanisms may be at work. At a broader scale I examined evolutionary transitions between yellow and anthocyanin pigmentation in the tribe Antirrhineae, and found that selection has acted repeatedly to promote the spread of pigment alleles. These results demonstrate that pollinators are a major determinant of reproductive success and mating patterns in wild *Antirrhinum*.

Contents

List of Tables

List of Figures

List of abbreviations

Chapter 1

s	Selection coefficient.
σ	Mean dispersal distance (distance between mother and offspring).
w	Cline width (inverse of the steepest gradient in a cline).
l	Characteristic scale of selection.

Chapter 2

d	Standard deviation parameter for density and frequency estimates.
s	Selection coefficient.
σ	Mean dispersal distance (distance between mother and offspring).
w	Cline width (inverse of the steepest gradient in a cline).
l	Characteristic scale of selection.
GLMM	Generalised linear mixed model.
AIC	Akaike's Information Criterion.
r	Rate of polliator visitation.

Chapter 3

FSG	Full-sibling group
MCMC	Monte Carlo Markov chain.
O	Set of individuals sharing a single mother, but with one or more fathers (a half sibling array).
F	Set of candidate fathers for individuals in O.
$\mathbf{G}$	Matrix of posterior probabilities that each male F is the father of each offspring in O.
g_{ik}	Posterior probability that the k^{th} male is the father of the i_{th} offspring; ik^{th} element of $\mathbf{G}$.
x	Proposed FSG within O.

C	Proposed partition of O into one or more FSGs.
γ	Matrix of posterior probabilities that each male in F is the father of each FSG in O; analogue of $\mathbf{G}$ but for sibships instead of individuals.
$\Pr(\mathbf{G}\|C)$	Likelihood of partition C.
$\Pr(C)$	Prior probabilty of partition C.
$\Pr(C\|\mathbf{G})$	Posterior probability of partition C; $\Pr(\mathbf{G}\|C)$.
$\Pr(x_i)$	Prior probabilty of the i^{th} FSG in O.
p	Probability parameter of the geometric distribution.
$\Pr(N)$	Prior probabilty of the number of sibships.
UPGMA	Unweighted pair-group method algorithm.
D	Pairwise distance matrix among full sibships; absolute values of each pairwise likelihood that two individuals in O are full siblings.
$\Pr(C\|\mathbf{G})_{true}$	Posterior probability of the true partition.
$\Pr(C\|\mathbf{G})_{max}$	Posterior probability of the most probable partition found.
λ	Per-locus likelihood that a male is the true father of an offspring.
ϵ_1	Genotype error rate per haploid SNP.
ϵ_2	Genotype dropout rate per diploid SNP.

Chapter 4

SULF/+	*SULF/SULF* or *SULF/sulf* genotype.
a	Scale parameter of the generalised gaussian distribution.
b	Kurtosis parameter of the generalised gaussian distribution.
d_{ij}	Distance between individuals i and j.

Chapter 5

q_{ij}	Instantaneous transition rate from phenotype i to phenotype j.
w	White.
a	Anthocyanin-pigmented.
y	Yellow.
d	Double-pigmented.
λ	Instantaneous speciation rate.
μ	Instantaneous extinction rate.
MCMC	Monte Carlo Markov chain.
pp	Posterior probability.

Chapter 1

Introduction

Much of biology is concerned with understanding the enormous diversity of organismal forms observed in nature. In particular, evolutionary biology seeks to understand the processes which bring about new variation, and the forces which allow new forms to spread and establish. Charles Darwin's (1859) key contribution was to describe the phenomenon of natural selection as a mechanism to promote the spread of "favourable races", which was subsequently demonstrated to have a solid genetic basis by the modern evolutionary synthesis in the early 20th century (Fisher, 1930; Dobzhansky, 1937; Bowler, 1989). Nevertheless, the principle of natural selection presents a paradox; if selection reduces variation by favouring only the fittest genotypes, how is it that we still observe so much variation in nature? To address this problem, we can break the question down into several components: which traits and genes are important for natural selection, what are the biotic and and abiotic agents which mediate it, and how does the influence of selection and genetic drift vary in time and space?

The evolutionary process can be observed at multiple levels. At the smallest level, alterations to DNA sequences cause structural or regulatory changes to molecules in cells. This changes the phenotype of the organism, influencing how the organism interacts with its environment, and how likely it is to survive to breed and pass on its genetic material. The cumulative effects of this variation across many individuals lead to genetic differentiation among populations or through time, and in some cases to complete reproductive isolation between species (Hart, 2011). Recent technical developments have allowed major insights to be gained, especially in the molecular basis of variation, but an integrated understanding of evolutionary processes

at different levels remains challenging for most systems. For example, studies on the molecular and developmental basis of wing patterning in the tropical butterfly *Heliconius* have shown that a relatively small toolkit of genes is sufficient to explain much of the diversity observed (Joron et al., 2006), but it is difficult to directly test the significance of this variation, nor the selective agents which cause it (Mallet and Barton, 1989b). In contrast, the pedigreed Soay sheep population on the St. Kilda archipelago show abundant inter-individual variation which correlates with parasite-load and climatic variables (Coltman et al., 1999; Coulson et al., 2001). It is nevertheless impractical and/or unethical to perform manipulative experiments to identify causal agents of selection, and furthermore the population is too small and the traits in question likely too polygenic to identify the underlying molecular variation. The application of modern methods to systems which can be examined at the molecular, ecological and population genetic levels is therefore likely to provide new insights into the evolutionary processes which have acted on them.

Natural environments are clearly much more complex than laboratory conditions, and organisms can behave surprisingly differently under field conditions. For example, growth chamber studies have identified several biochemical pathways involved in flowering time regulation in the model plant *Arabidopsis thaliana*, but a host of other biotic and abiotic stressors can influence phenology in the wild (reviewed by Anderson et al., 2014). It is therefore no surprise that loci showing an association with flowering time in wild populations do not overlap with those identified in growth chamber studies (Weinig et al., 2002; Brachi et al., 2010; Dittmar et al., 2014). For long-lived organisms, it is also important to collect data over many years to capture the lifetime reproductive success of the organism. Moreover, environmental variation between years is often substantial, which means that a single season may not be representative of broader trends. This is beautifully illustrated by Grant & Grant's (2002) 30-year study on Darwin's finches on the Galápagos islands, which demonstrated that periodic droughts caused changes in seed size, which in turn led to higher survival among birds with large beaks. These examples highlight the need to examine the significance of biological variation in its natural context, as well as the advantages gained through long-term studies.

1.1 Hybrid zones

Hybrid zones occur where genetically diverged populations abut, and hybrid individuals are formed. In most cases (Barton and Hewitt, 1985) hybrid zones reflect a balance between the dispersal (σ) by the parental types, and selection (s) either against heterozygotes and recombinants (Haldane, 1948; Key, 1968; Slatkin, 1973). Hybrid zones are characterised by the formation of a sigmoidal gradient (cline) in allele frequency, or the mean of a quantitative trait defined by the width w, the inverse of the steepest part of the cline. Narrow clines imply strong selection. Maintenance of hybrid zones in migration-selection balance requires that w be greater than the characteristic scale of selection $l = \sigma/\sqrt{s}$ (Slatkin, 1973). Such hybrid zones can be very long, and remarkably robust in width along their entire length. For example, the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata* runs for some 4000km from Poland to Croatia, but maintains a width of around 6km across most of its range (Szymura and Barton, 1991).

Hybrid zones have a number of attractive properties for the investigation of selective processes in the wild. Firstly, by rearranging the equation for l and estimating cline width and dispersal, we can get a direct estimate of the selection strength maintaining the cline for a given locus. Dispersal can be estimated directly, or inferred from linkage disequilibrium among clines (Barton and Gale, 1993). This estimate of selection strength has two advantages. Firstly, it represents the cumulative effect of selection over many generations, and therefore averages over inter-annual variation. Secondly, it represents the total selection acting on all aspects of the phenotype in which the locus is involved, including direct effects, pleiotropy, and effects of linked loci. Finally, hybrid genotypes bring together combinations of alleles which have never previously been tested by selection, and therefore represent evolutionary laboratories for novel variation (Hewitt, 1988). Patterns of allele frequencies are therefore an invaluable tool in the investigation of the genetic factors and selective forces which contribute to population differences.

1.2 The evolution of flower colour

There are number of aspects of both plant and human biology which make flowers attractive to both pollinators and evolutionary biologists alike. Firstly, variation in flower colour is conspicuous, and rare mutations are often pre-

served by horticulturalists. Secondly, human fascination with flower colours is sufficient to make cut flowers big business, with annual sales of $70 billion in 2006 (Grotewold, 2006). From the research this has stimulated (Tanaka et al., 1998) we now have an excellent understanding of the biochemical pathways involved in pigment synthesis, and the underlying genetic regulatory network (reviewed by Grotewold, 2006). The best understood pathway is that of the anthocyanins, a class of water soluble flavonoids derived from chalcone. After the synthesis of naringenin, the pathway forks into three sub-pathways which result in the blue delphinidin, red/orange pelarganoidin and pink cyanidin pigments. The majority of yellow flowers are derived from carotenoid, or more rarely aurone pigments, controlled by separate biochemical and genetic pathways (Harborne, 1966; Nakayama, 2002). Based on this, several studies on the molecular basis of evolutionary transitions in flower colour have shown that evolution can be remarkably repeatable. Firstly, although mutations are just as likely to hit biosynthetic as regulatory genes, the pleiotropic effects of structural genes across tissues means that only tissue-specific mutations at regulatory genes are able to fix in natural populations (Streisfeld and Rausher, 2011). Furthermore, transitions tend to be due to single loss-of-function mutations, and show a strong association with a handful of transcription factors (Whittall et al., 2006; Rausher, 2008; Smith and Rausher, 2011; Streisfeld and Rausher, 2011; Wessinger and Rausher, 2013). Floral pigmentation has therefore yielded unique insights into the nature of evolutionary transitions which would otherwise be extremely difficult to elucidate.

Moreover, floral pigmentation is strongly linked to plant fitness. Animal-pollinated plants rely on pollinators to disperse their gametes, and colour is among the most important signals plants use to recruit pollinators. Different colours show associations with different pollinator groups (Faegri and Van der Pijl, 1966), and variation in flower colour has well-documented effects on pollinator attraction and hence reproductive success (e.g. Waser and Price, 1981, 1983; Stanton et al., 1986; Stanton, 1987a; Bradshaw and Schemske, 2003; Hopkins and Rausher, 2012). There is also a growing recognition that floral pigments are involved in a host of non-pollinator related traits (Strauss and Whittall, 2006), such as drought tolerance (Warren and Mackenzie, 2001), herbivore resistance (Irwin et al., 2003), and heat stress (Coberley and Rausher, 2003). Because many of the species in question are annual herbs, they are amenable to a combination of field studies and manipulative experiments to isolate components of selection (Mitchell-Olds and

Shaw, 1987). Furthermore, pigmentation polymorphism often has a simple genetic basis, which means that it is often possible to identify causal loci. In a classic study, Bradshaw and Schemske (2003) repeatedly backcrossed alleles at the *YELLOW UPPER* locus into the pink-flowered *Mimulus lewisii* and the red-flowered *M. cardinalis* and showed that the substitution at a single locus was sufficient to effect a dramatic shift in visitation by bee and hummingbird pollinators. Flower colour polymorphisms therefore have great potential for the investigations of the genetic basis of adaptation, the agents of selection, and the effects on reproductive success in wild populations.

1.3 *Antirrhinum* as a model system

Antirrhinum is a genus of around 20 annual or short-lived perennial herbs in the tribe Antirrhineae, which are found mostly in the Iberian peninsula (Rothmaler, 1956; Stubbe, 1966). *Antirrhinum* thrives in disturbed habitats, and is especially common along roadside and railway embankments. All species except the domesticated *A. siculum* have a gametophytic self-incompatibility system (Xue et al., 1996), although self-compatible mutants are occasionally seen in the wild at very low frequency. The flowers have five petals fused into a tube, with the three ventral petals forming a characteristic hinged mouth structure, which pollinators must open to access nectar (figure 1.1). Considerable strength and dexterity is needed to achieve this, and as such almost all pollination services are performed by large bees, especially *Bombus* and *Xylocopa* species (Vargas et al., 2010). Such is the extent of specialisation for pollination by large bees that Proctor and Yeo (1973) described *Antirrhinum* as among the best examples of the 'bee pollination syndrome', owing to its bee-conspicuous colours, zygomorphic form and landing platform.

Antirrhinum, and *A. majus* in particular, has a long history as a study organism for geneticists (Schwarz-Sommer et al., 2003). Indeed, its use as a model organism came about amidst the fierce debate at about the discrete or continuous nature of biological variation at the turn of the 20th century following the rediscovery of Mendel's work, mirrored by a similarly intense controversy at the University of Cambridge about the role of women such as Muriel Wheldale who pioneered early genetic research under William Bateson (Bowler, 1989; Richmond, 2001). *Antirrhinum* was adopted as a model system by classical geneticists working in the UK and Germany

Figure 1.1: Bumblebee entering an *Antirrhinum* flower, showing the hinged mouth structure. Photograph: Marie Mirouze.

because of the abundant, conspicuous and heritable variation in floral phenotypes. The ease of cultivation and crossing meant that *Antirrhinum* made key contributions to early understanding of inheritance, biochemistry and genetic linkage (Wheldale, 1907; Baur, 1911; Onslow (née Wheldale), 1916; Baur, 1924; Stubbe, 1966). Antirrhinum was later to be at the forefront of early work on transposons and transposon tagging for mapping (Wienand et al., 1982; Coen et al., 1986), and has also provided key insights into how regulatory signals can organise cells into complex floral forms (Green et al., 2010). Furthermore the genome is of modest size (∼700MB) and diploid, and full sequencing is currently underway. These developments show that *Antirrhinum* is just as pertinent a model system now as it was a century ago.

A key aim of early genetic work was to establish the hereditary basis of flower colour. Among the most important genes identified were those controlling the distribution of pigments in the face of the flower where the ventral lips curl forward. The *ROSEA* (*ROS*) locus controls the distribution of one or more anthocyanins, and in the absence of yellow pigmentation homozygotes are either solid magenta (*ROS/ROS*), or white (*ros/ros*). Heterozygotes can be distinguished by the less intense pink colour, and by the patchy distribution caused by the tightly linked *ELUTA* (*EL*) locus (Baur, 1924). *Rosea* is actually a tightly linked cluster of three *MYB*-transcription factors, but much of the variance in anthocyanin expression is explained by a 150-bp deletion upstream of *ROS1* (Schwinn et al., 2006). The *Sulfurea* (*SULF*) locus controls yellow aurone pigmentation throughout the face (Wheldale, 1907; Baur, 1924). The unpigmented allele is fully dominant over the pigmented allele, but the molecular basis of *SULF* has yet to be elucidated. F_2 crosses between *ROS SULF/ROS SULF* and *ros sulf/ros sulf* lines give rise to six distinct phenotypes (figure 1.2), including unpigmented white phenotypes, and double pigmented orange phenotypes. The loci identified in mutant lines match the loci segregating in the wild (Whibley et al., 2006), and crosses among species demonstrate that control by *ROS* and *SULF* of floral pigmentation is shared across the genus (Hackbarth et al., 1942). Given the ecological relevance of pigments noted above, the independent control of pigment classes represent a fascinating opportunity to investigate the interaction among loci in the wild.

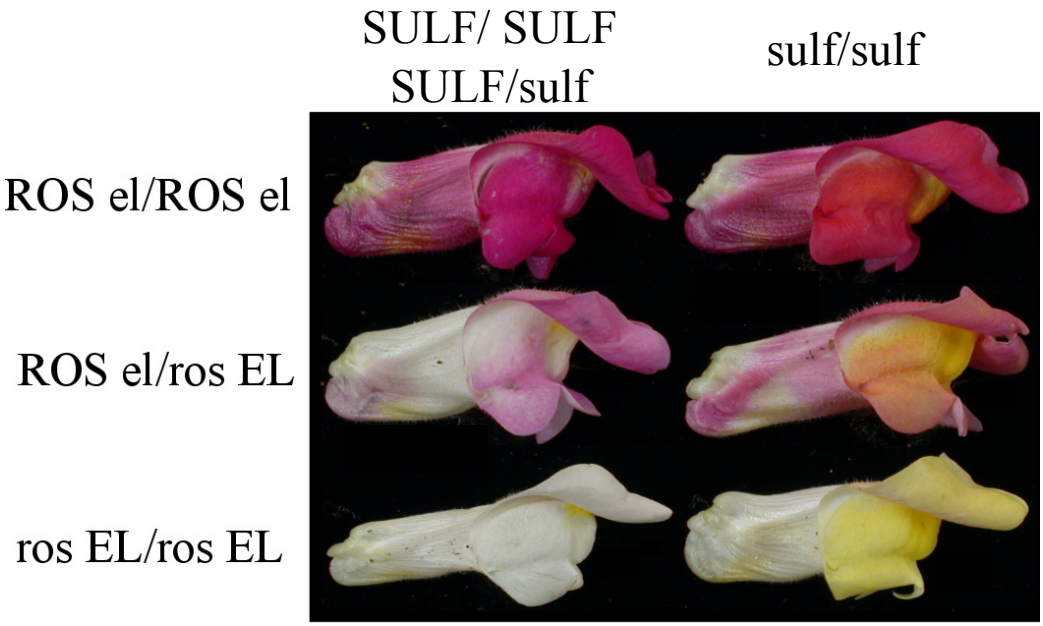


Figure 1.2: Combinations of alleles at the *Rosea* and *Sulfurea* loci controlling anthocyanin and aurone pigmentation give rise to six flower phenotypes. The *SULF* allele is fully dominant over *sulf*, so *SULFUREA* heterozygotes have no yellow pigmentation in the face. *A. m.* pseudomajus shows the magenta phenotype in the top right, and *A. m. striatum* shows the yellow phenotype in the bottom right. Note that the yellow patch at the dorsal tip of the ventral petals is present in almost all plants and likely under separate genetic control.

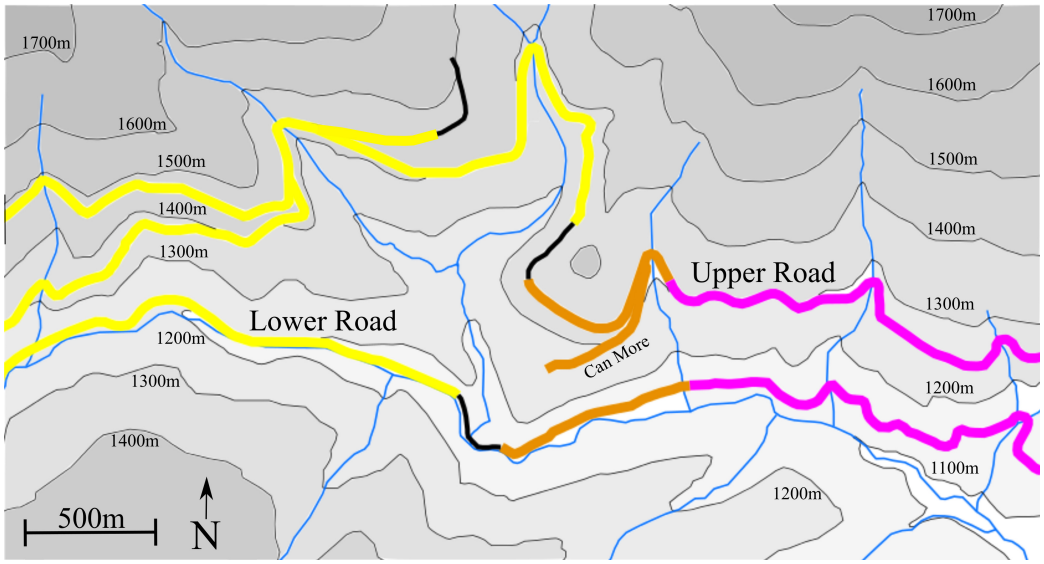


Figure 1.3: The Vall de Ribès hybrid zone. *A. majus* grows along two roads (upper and lower) labelled yellow or magenta for predominantly *A. m. striatum* and *A. m. pseudomajus* respectively. Orange indicates the core hybrid zone, and thick black lines roads where *Antirrhinum* are not observed. Blue lines are streams, and thin lines indicate 100m contour boundaries.

1.3.1 The *Antirrhinum majus* hybrid zone

The discovery of a natural *A. majus* hybrid zone has greatly facilitated investigation into the evolutionary dynamics of this species in the wild. The magenta-flowered *A. m. pseudomajus* and the yellow-flowered *A. m. striatum* are distributed parapatrically throughout the Pyrenees mountains, and form a hybrid zone where their ranges meet along two roughly parallel roads in the Vall de Ribès, approximately 120km north of Barcelona (figure 1.3). Here, hybrids are confined to a narrow stretch of ∼1km, and sharp clines for colour phenotypes and *ROS* alleles indicate that colour is under strong selection (Whibley et al., 2006). There is no evidence of genetic incompatibilities (Andalo et al., 2010), nor of any differences in nectar reward among morphs (Tastard, 2009). Furthermore, there is no obvious change in floral or pollinator community, and 90% of pollination is carried out by a single species of bumblebee, *Bombus hortorum* (Rebecca Franklin, unpublished data). This raises the intriguing question of what mechanisms maintain the stability of

the hybrid zone year upon year.

To elucidate the evolutionary processes at work, this population has been subject to a collaborative project between groups at IST Austria, The John Innes Centre (Norwich, UK) and the University of Toulouse. One focus has been to characterise major molecular variants segregating in the population, especially *EL* and *SULF* (Whibley, 2004; Tavares, 2014). Another aims to collect the locations, phenotypes and tissue from the entire hybrid zone population across multiple years, and to build this into a pedigree, providing direct measures of variance in reproductive success and the factors which contribute to it. Finally, an analysis of the broader scale patterns of cline shape and population structure aims to elucidate gene flow and selection at different spatial scales. A combination of approaches such as this offers a unique insight into the evolutionary process from the small to large scales.

1.4 Thesis aims

Despite this, direct evidence is still needed for the agents of selection. Given the diverse roles of pigment molecules in plant biology (Strauss and Whittall, 2006) these may be numerous. Nevertheless, the link between flower colour, pollinator behaviour and plant mating success strongly implicates pollinators as potential selective agents. In this thesis I explore the evolutionary history of interacting pigments, and investigate the role of pollinators in shaping the patterns we see today.

The specific aims of this thesis are to:

1. establish whether pollinators forage non-randomly in the hybrid zone to establish whether they could act as agents of selection;

2. identify the traits which influence pollinator visitation;

3. quantify the effect of pollinator behaviour on the reproductive success of *Antirrhinum;*

4. examine the patterns of flower colour evolution over longer time scales to assess the historical role of selection and genetic constraint on extant patterns of diversity observed in nature.

As part of aim three it will also be necessary to assess whether genotype data from wild-pollinated seedlings can be used to directly infer mating patterns

in the wild, and to develop methods for doing so. To achieve these aims this thesis uses a range of field, experimental, genealogical and phylogenetic data to demonstrate link between pollinators, reproductive success and long-term genetic changes at as many levels as possible.

Chapter 2

Pollinator behaviour causes assortative mating and frequency-dependent pollinator selection for flower colour in a snapdragon (*Antirrhinum*) hybrid zone

2.1 Abstract

The fitness of an individual can depend not only on its own genotype, but also on the genotype of its neighbours. Selection against rare phenotypes can maintain sharp hybrid zones, and assortative mating can strengthen the consequent reproductive isolation. This is crucial for flowering plants that rely on pollinators for dispersing male gametes (i.e. pollen), particularly when foraging decisions are based on their local phenotypic context. Here, we examine the strength and mechanisms of pollinator mediated selection on flower colour variation between two subspecies of the snapdragon *Antirrhinum majus*. Where there is gene flow between the yellow- and magenta-flowered subspecies, narrow hybrid zones form, generating novel colour phenotypes due to recombination between major pigmentation genes. We used a novel tagging method to infer rates of pollinator visitation to wild plants in a natural hy-

brid zone, and found that plants receive more visits if their colour is more similar to their neighbours. Using experimental arrays and naïve pollinators, we manipulated the frequency of the parental phenotypes to examine pollinator behaviour and subsequent changes in allele frequencies of the progeny in the next generation. The pollinators showed strong assortative foraging, but no overall preference for colour or dependence on frequency. Pollinator behaviour was consistent with phenotype frequencies in the following generation. We found a significant deficit of hybrid offspring, supporting direct observations that the pollinators drive assortative mating. Our results provide direct evidence that pollinator-mediated selection on flower colour can generate strong reproductive isolation, and provides support for the interaction between assortative mating and positive frequency dependent selection as the underlying mechanism.

2.2 Introduction

As populations diverge they accumulate genetic differences which act as a barrier to interbreeding, or reduce the fitness of hybrid offspring. When these differences become so strong that populations can no longer interbreed at all, they split into distinct species (Mayr, 1942; Coyne et al., 2004; Hart, 2011). Such differences are thought to be caused by the fixation by drift or selection of alleles in each population which cause then negative epistatic interactions when brought together in hybrids (Orr, 1996). Examination of the specific mechanisms which underly reproductive isolating barriers can therefore shed light on the traits which are important for adaptation, and the roles of selective and neutral processes in shaping diversity among wild populations.

Biologists have long noted the utility of studying pigmentation variation in natural hybrid zones for understanding the evolution of reproductive isolating barriers. Hybrid zones occur where populations meet and interbreed, with hybrids restricted to a contact zone maintained by a balance between selection and dispersal (Haldane, 1948; Slatkin, 1973; Barton and Hewitt, 1985). Here, recombination breaks up associations among adaptive and non-adaptive loci and traits, giving rise to new combinations which have not been tested by selection elsewhere (Hewitt, 1988). Colour is an important signal to potential predators, mates, and mutualists, and it is unsurprising that nine of the 21 hybrid zones listed by Barton and Hewitt (1985) involve variation

for pigmentation (see also Haldane (1948), and floral hybrid zones listed below). Among angiosperms, variation for flower colour abounds in natural populations (Warren and Mackenzie, 2001; this volume, chapter 5) and has long been recognised as an important signal to attract pollinators (Darwin, 1876; Faegri and Van der Pijl, 1966). Moreover, there is growing recognition of the links between floral pigmentation and a host of physiological traits such as drought tolerance or herbivore resistance (Warren and Mackenzie, 2001; Strauss et al., 2004; Strauss and Whittall, 2006). This variation makes flower colour an intriguing testing ground for hypotheses about the evolution of reproductive barriers in the wild.

Flower colour variation in hybrid zones have been extensively studied in a number of systems, such as *Aquilegia, Ipomopsis, Iris* and *Mimulus* (Fulton and Hodges, 1999; Campbell et al., 1997; Emms and Arnold, 2000; Streisfeld et al., 2013). However, these studies compare full species pairs, which differ not only for flower colour, but many other floral and ecological traits (Hodges and Arnold, 1994; Emms and Arnold, 1997; Campbell, 2004), as well as major pollinator phyla (typically hummingbirds *vs.* moths or bees). This makes it difficult to distinguish the traits that were historically important for divergence from those which evolved secondarily. Studies of flower colour polymorphisms within a species have focussed on mixed populations, and have frequently identified pollinator discrimination against particular flower colour morphs (e.g. Waser and Price, 1981; Brown and Clegg, 1984; Stanton, 1987a). Such studies are instructive as to the mechanisms which maintain variation within a population, but tell us little about the mechanisms maintaining reproductive isolation between species (but see Hopkins and Rausher, 2012). Furthermore, in all but a handful of cases (Waser and Price, 1981; Brown and Clegg, 1984; Tastard et al., 2012), field studies of floral colour variation have used plants in experimental arrays, which may not accurately reflect the environment experienced in the field. Studies examining isolating barriers between populations at the very early stages of divergence are therefore lacking.

Animal foragers often restrict feeding to a subset of available items, even when other palatable or rewarding items are available. In plant-pollinator systems, this has been most commonly studied in the context of pollinator constancy, where a pollinator visits the same species or floral morph sequentially (e.g. Darwin, 1876; Grant, 1950; Mather, 1947; Bateman, 1951; Brown and Clegg, 1984; Stanton et al., 1989). In contrast, studies of predator-prey systems have focussed on frequency-dependent selection, where predators

favour rare species or morphs by preferentially predating on locally common morphs (Clarke, 1962; Allen and Clarke, 1968; Allen and Greenwood, 1988). Frequency-dependent foraging by pollinators has received much less attention (reviewed in Smithson, 2001). Pollinators readily exhibit positive frequency-dependent foraging behaviour when foraging on rewarding artificial inflorescences (Smithson and Macnair, 1996, 1997). Experimental arrays have sometimes found the same behaviour under field conditions (Epperson and Clegg, 1987; Eckhart et al., 2006; but for counter-examples see Waser and Price, 1981; Jones, 1996b), although the effect is much weaker than observed in the laboratory (Smithson, 2001). Levin (1972) and Epperson and Clegg (1987) demonstrated reduced outcrossing rates for rare morphs, but to our knowledge only Tastard et al. (2012) have found significant positive correlations between frequency and seed set (cf. Levin and Kerster 1970; Waser and Price 1981; Jones 1996a; Smithson 2001). The proximate reason for such selective behaviour is likely to be that there is a net gain in foraging efficiency by focussing on a subset of prey or flowers (Allen and Greenwood, 1988). For example, selective foraging may increase efficiency if certain flowers are easier to locate in the current environment, pollinators have experience with handling particular flowers and know them to be rewarding, or if reduced handling time offsets increased travelling time (Waser, 1986; Chittka et al., 1999; Spaethe et al., 2001). The dependence on such factors on the local environment highlights the importance of context in plant-pollinator relationships.

Frequency-dependent foraging and floral constancy have distinct effects on plant reproduction. When flowers are rewarding, pollinator preference for common species or morphs will exert positive frequency-dependent selection on the population, driving the rare form to extinction, and placing a constraint on floral diversity (Smithson and Macnair, 1996). Meanwhile, floral constancy will facilitate pollen transfer between plants with similar floral phenotypes, leading to assortative mating (Darwin, 1876; Mather, 1947). In the context of a hybrid zone, this means that frequency-dependent foraging will directly contribute to the maintenance of clines (Mallet and Barton, 1989a), whilst constancy will strengthen the barrier to gene flow caused by those clines. Most studies of frequency dependence and assortative mating in plants have considered only one of these phenomena, but they must to some extent be interdependent. On one hand, preference for common flowers will lead to sequential visits to those flowers, and hence assortative mating. Similarly, if pollinators show constancy to a morph they have previ-

ously visited, but choose that morph at random, they are more likely to pick the frequent morph just by chance, leading to positive frequency-dependent selection. A stark example of assortment causing frequency-dependent selection is provided by *Partula* snails, where assortative mating for coil chirality alone is sufficient to effect strong frequency-dependent selection, leading to a population mosaic of locally fixed morphs (Johnson, 1982). This interaction between frequency-dependent selection and assortative mating demonstrates that the two phenomena cannot be considered in isolation.

In this study we examine a hybrid zone between two subspecies of the snapdragon *Antirrhinum majus* which differ at genes controlling yellow and magenta floral pigmentation. The yellow-flowered *A. m. striatum* and the magenta-flowered *A. m. pseudomajus* are distributed parapatrically throughout the Eastern Pyrenees and form narrow hybrid zones where their ranges meet. Here, alleles at pigmentation loci recombine to generate striking transgressive variation in flower colour, including white, pink and orange phenotypes (Whibley et al., 2006; Tastard et al., 2008). This sharp transition in phenotype and allele frequencies implies strong divergent selection for flower colour on either side of the hybrid zone (Whibley et al., 2006). There is no evidence for postzygotic barriers (Andalo et al., 2010), yet *Antirrhinum* in this population are pollinated almost exclusively by a single pollinator, *Bombus hortorum* (Rebecca Franklin and Andrew Bourke, unpublished data). Jaworski et al. (2015) showed that the closely related *B. terrestris* has no innate preference for either subspecies. If pollinator behaviour does indeed contribute to the maintenance of this hybrid zone, this therefore requires that the same pollinators display a flexible, context-dependent preference of the kind described above. Tastard et al. (2012) demonstrated that seed set is positively correlated with how similar plants are to their neighbours, but could not directly ascribe this correlation to the action of pollinators. This system therefore offers a promising opportunity to investigate the roles of frequency-dependent foraging and floral constancy in pollinator behaviour, and their influence on the maintenance of the *Antirrhinum* hybrid zone.

We investigated pollinator-mediated selection on flower colour in the *Antirrhinum* hybrid zone using a combination of field and experimental assays. We first used an indirect tagging approach and an intensive, large-scale demographic survey of this population to assess patterns of pollinator visitation among wild plants over multiple years. However, wild plants have a patchy spatial distribution with clustering of particular phenotypes across the hybrid zone, whilst wild pollinators may be influenced by their experience of

the broader floral community. To better account for these factors we also performed controlled array experiments to investigate patterns of frequency-dependent foraging and constancy by näive pollinators. Our results point to strong differences in context-specific behaviour by pollinators, causing strong selection on pigmentation loci across the hybrid zone.

2.3 Materials and Methods

2.3.1 Study population

Antirrhinum majus striatum and *A. m. pseudomajus* form a ∼1km contact zone along two roads (upper and lower) in the Vall de Ribès at 1150m and 1300-1400m respectively (figure 1.3) . *Antirrhinum* flowers have a characteristic hinged mouth structure which pollinators must open to access nectar. The bumblebee *Bombus hortorum* performs more than 90% of pollination service across the hybrid zone (Rebecca Franklin and Andrew Bourke, unpublished data). Previous studies found no detectable genetic incompatibilities between the subspecies (Andalo et al., 2010), nor any difference in nectar production in the hybrid zone (Tastard, 2009).

Flower colour can be classified into two yellow and three magenta phenotype classes based on the scoring system developed by Whibley et al. (2006). The distribution of magenta anthocyanin pigments is controlled by the *ROSEA* (*ROS*) and *ELUTA* (*EL*) gene complex (Schwinn et al., 2006; Baur, 1924). *ROSEA* and *ELUTA* are tightly linked, and recombinants are rare in the wild (Tavares, 2014). *ROS el/ROS el* plants have a smooth distribution of anthocyanins throughout the ventral lip, whilst *ros EL/ros EL* plants have no anthocyanin in the ventral lip. *ROS el/ros EL* heterozygotes typically have an irregular patchy distribution of anthocyanin in the ventral lip caused by interference by *ELUTA*. The *SULFUREA* (*SULF*) locus controls the distribution of yellow aurone pigmentation in the face. The *SULF* allele is fully dominant over *sulf*; *sulf/sulf* plants are yellow throughout the face, while *SULF/SULF* and *SULF/sulf* plants have only a patch of yellow at the centre of the ventral petals. Based on anthocyanin and aurone pigmentation, plants can be categorised into one of six phenotype classes (figure 1.2). Flowers do not reflect strongly in the UV spectra (Tastard et al., 2008).

We have collected demographic data on this population over multiple years. In each year, we recorded the GPS position and flower number of

most of the plants within 2km of the core of the hybrid zone using a Trimble GeoXT (n=2299, 1984 and 2228 in 2010, 2011 and 2012). We also collected a flower from each plant, which we visually scored for pigmentation phenotype to infer the genotype for *ROSEA* and *SULFUREA*.

2.3.2 Visitation rate in wild plants

Insects are only rarely directly observed pollinating *Antirrhinum* in the hybrid zone, which precludes direct observations of visitation rates. Instead we employed an indirect tagging method to infer pollinator visitation for a greater sample of plants. We placed 5mm circular transparent cellophane tags into the side of the "mouth" of the flower such that the tag would not impede pollinator access. Pressure from the dorsal petals holds the tag in place until a pollinator opens the flower and the tag drops out. Assays were performed on 71, 76 and 54 plants on the lower road in 2010, 2011 and 2012 respectively and on 54 plants on the upper road in 2011. We aimed to choose roughly equal numbers of each phenotype, and all plants were located in the core of the hybrid zone. We tagged two healthy flowers on each of up to two inflorescences per plant, but explicitly account for this in statistical analyses. We labelled tagged plants with red jewellery tags which are inconspicuous to bees and marked tagged flowers with a dot of ink at the base of the calyx to aid re-identification. We then performed two surveys per day at 0900-1100 and 1600-1800, performing 12, 7 and 9 surveys on the lower road in 2010, 2011 and 2012 respectively, and 8 surveys on the upper road. At each survey we checked each flower for the presence or absence of a tag. If a tag was missing we inferred that a pollinator had visited the flower. We replaced any missing tags, or any which had fallen inside the corolla.

We used the demographic survey data to characterise the local floral neighbourhood for each plant. To estimate the density of neighbours around each tagged plant we applied a gaussian function to the euclidean distances between each focal plant and all other plants in the population, and summed across the resulting vector. We then used these densities to estimate the local phenotypic frequency of each plant as the density of neighbours of the same phenotype class relative to the density of all neighbours. To determine the most appropriate standard deviation, δ, of the gaussian function we estimated density and phenotypic frequency for all integer values of δ between 1m and 100m, and refitted the GLMM with all nine parameters (see below) for each δ. The lowest deviance was found with using $\delta = 3$m, but this distance is on

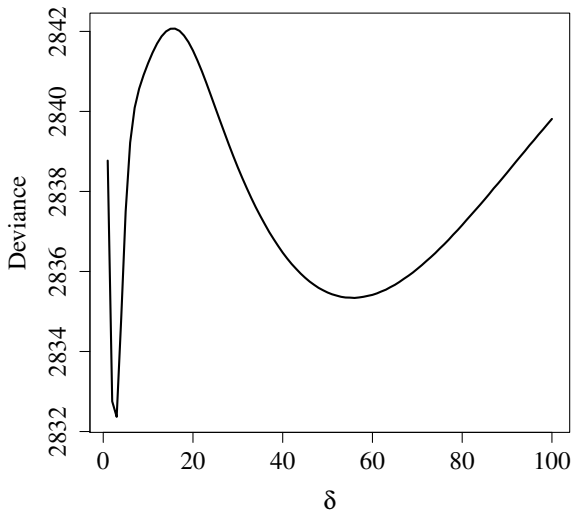


Figure 2.1: Deviance of the full nine-parameter model using estimates of density and frequency based on gaussian functions with different standard deviations δ.

the same order as GPS errors. We instead chose δ=56m because this gave a model with the next lowest deviance, and is sufficiently broad to be largely insensitive to GPS measurement error (figure 2.1).

To determine which variables contribute to pollination success in the wild we fitted a generalised linear mixed model (GLMM) to the tag data. We modelled data as a binary outcome at the individual level, i.e. whether a plant had received a visit to any flower at a given survey. We began by fitting a saturated model with 9 parameters, including 6 main effects: (1) local snapdragon density, (2) phenotypic frequency, (3) phenotype class, (4) number of flowers per plant, (5) upper vs. lower road, and (6) the number of flowers tagged per plant. The dataset is unlikely to have sufficient power to assess all possible interactions, so we also fitted only pairwise interaction terms between frequency, density and phenotype. We used a random effect for plant identity and a nested random effect of survey within year to account for correlations between readings from the same plant, and temporal correlations. We then proceeded to remove fixed effects which accounted for the least amount of variation based on χ^2 test comparisons of Akaike's

Information Criteria (AIC) (Zuur et al., 2009) to arrive at the minimum model required to explain the data. Analyses were performed in R using the lme4 package assuming a binomial error structure (Bates et al., 2011; R Core Team, 2012).

The strength of selection acting on a trait depends on effect of that trait on relative fitness, ω. The relationship between visitation rate r and plant reproductive success will be a complex function of resource availability, pollen limitation and flower number, which we do not as yet have information on. We therefore assume that flower number and resource availability are random with respect to traits and examine the cases for complete ($\omega = {}^{r}/_{2\bar{r}}$) and non-existent ($\omega = {}^{r}/_{\bar{r}}$) pollen limitation, representing upper and lower bounds for selection. Using these estimates of ω, we then estimated the strength of selection acting trait z as the covariance between ω and z (Lande and Arnold, 1983).

False positive errors occur when a tag drops even though no pollinator has visited, while false negatives occur when a tag does not drop even when a pollinator does visit. To quantify the frequency of false positives we placed tags in 84 flowers on 24 plants in July 2014 and placed bags around the inflorescence to exclude pollinators. We surveyed these plants for missing tags twice a day for eight days, or until no flowers remained. To quantify the frequency of false negatives we placed orange paper tags in all flowers of six plants and observed 171 pollinator visits directly and recorded whether the tag dropped or not.

2.3.3 Experimental arrays

The main pollinator species in the wild is not available commercially, so we instead acquired 15 hives of the congener *Bombus terrestris* from a breeder (Koppert, Netherlands). Whilst *B. terrestris* will not exactly match the behaviour of *B. hortorum,* we feel it is a useful model for two reasons. Firstly, the two species closely related and have similar sensory systems. Secondly, although wild, *B. terrestris* tend to rob nectar from *Antirrhinum* by chewing a hole through the base of the corolla rather than enter the flower directly (Christophe Andalo and Monique Burrus, unpublished data), robbing was rare in our experiment, and most pollinators entered the flower to access rewards. Thus, the *B. terrestris* in the experiment behaved more like wild *B. hortorum* than wild *B. terrestris*.

We grew seeds from 48 second generation families from three magenta

A. m. pseudomajus and three yellow *A. m. striatum* populations collected far from the hybrid zone. Sampling and crosses are described in detail by Jaworski *et al.* (2015; populations used are the Pra, Bes, Cal, Thu, Lys, and Vil populations described in that paper). Due to constraints on time and space we did not include any hybrid phenotypes. We soaked approximately 100 seeds of each family in 0.3ml millipore water for 48 hours at 4^{o}C, before sowing them on filter paper in petri dishes. We moved petri dishes to a growth chamber on a 16 hour cycle under Phillips GreenPower LED lamps. Once cotyledons had emerged we transferred seedlings to 9cm pots filled with potting compost (Gramaflor) in a greenhouse with no additional lighting.

We set up arrays of 36 plants in a 3x3x2.5m organza tent to exclude native pollinators. We arranged plants in an offset grid so that each plant was 40cm from its nearest neighbours in all directions, with a single bumblebee hive in the south-west corner of the tent. We applied five frequency treatments of yellow to magenta flowers: 6:30, 12:24, 18:18, 24:12 and 30:6. We performed a total of three replicates of each treatment in August 2013, September 2013 and May 2014. The bees had no previous experience of handling flowers, and we used each colony only for a single frequency treatment and replicate. We allowed bees to forage in the tent for 24 hours to learn how to handle *Antirrhinum* flowers, and then observed each individual bee foraging for three periods of 1.5 hours between 0900 and 1700. We observed individual bees for ten minutes or until she returned to the hive, and recorded the plants and flower positions she visited. We classified transitions between plants as yellow-yellow (YY), yellow-magenta (YM), magenta-yellow (MY) or magenta-magenta (MM). We further grouped these into heterospecific (YM and MY) and conspecific (YY and MM) transitions

We applied two GLMMs to data on pollinator transitions between plants using the R package lme4 (Bates et al., 2011; R Core Team, 2012). To assess the degree of frequency- or colour-dependent preference we modelled transitions to yellow plants as a function of the frequency of yellow plants in the treatment and whether the transition was hetero- or conspecific. Note that the choice of yellow as a reference is arbitrary. To assess the degree of assortative foraging we modelled the probability of a conspecific transition as a function of the frequency of the target plant, and whether the target was yellow. We accounted for variation among replicates and individual bees by fitting these as random effects. We performed analyses on only those transitions from plants which were not on the edge of the array to account for edge effects. To confirm that the observed results were really due to the

effect of flower colour we randomly permuted the colours of plants in each array 10,000 times. We then refitted both GLMMs to each permuted dataset and compared the observed estimates of the intercept and regression coefficients to their distributions for permuted datasets. If observed regression coefficients differ from those for permuted data, this indicates a frequency-dependent effect on transitions, whilst a difference in intercepts indicates an overall preference for transitions in a particular direction. This permutation approach allowed us to assess the effect of flower colour whilst holding other variables constant.

2.3.4 Offspring phenotype frequencies

We moved plants from the experimental arrays to a pollinator-excluding greenhouse to allow seeds to develop. Plants from the September 2013 replicate set very few seeds owing to record temperatures during fruit development, so we excluded this replicate from further analyses. We germinated and transplanted seeds from the remaining two replicates as described above and allowed plants to mature either outside in Summer 2014 or inside the greenhouse in Winter 2014-15 under Sylvania Grolux lights on a 16:8-hour light:dark cycle.

Once plants had flowered we scored them as magenta, yellow or hybrids based on *ROSEA* phenotypes. Since the arrays contained only yellow and magenta parents, yellow or magenta progeny must have resulted from a conspecific mating event, whilst hybrid offspring must be the result of an interspecific mating event. We then compared the observed frequencies of *ROS* alleles and hybrid phenotypes to expected frequencies given Hardy-Weinberg using χ^2 tests with one degree of freedom. To allow a comparison of the strength of assortative mating inferred from offspring phenotype frequencies and pollinator observations we also calculated Sobel and Chen's (2014) index of reproductive isolation:

$$RI = 1 - 2(\frac{H}{H + C})$$

where H and C denote the numbers of heterospecific and conspecific transitions or mating events. This is algebraically equivalent to Bateman's Index of floral constancy (1951) and ranges from 1 for complete assortment to -1 for complete disassortment.

2.4 Results

2.4.1 Wild pollinators prefer common morphs

To assess patterns of pollinator visits to wild plants we placed tags inside flowers which fall out when a pollinator enters. We made a total of 2449 observations on plants in 38 surveys, for which we inferred 954 pollinator visits. We found a significant, but extremely weak, negative effect of the number of flowers on pollinator visitation, as well as a weak effect of the number of flowers tagged (table 2.1). By far the most important variable that explained the observed variation in missing tags is the local phenotypic frequency of a plant. An increase in frequency of 10% led to a 7.4% increase in the probability of a pollinator visits per survey (figure 2.2). This corresponds to a selection coefficient for phenotypic frequency of $s = 0.896$ (SE=0.132) in the case of no pollen limitation, although this may be as much as double this as pollen limitation increases.

We performed 558 surveys of bagged flowers, and observed 12 tags (2.2%) to have dropped in the absence of pollinators. Six of these were during hail storms in 2014, which never occurred in 2010-2012. We directly observed 171 pollinator visits to tagged flowers and found that the tag failed to drop in 22 out of 179 visits (12.9%).

2.4.2 Naïve pollinators forage assortatively

We observed 2363 pollinator transitions between plants during 401 foraging bouts in controlled arrays, and compared observed patterns of pollinator movements to permuted datasets (table 2.2). Across treatments, transitions to yellow were around 60% more likely if the previous flower visited was also yellow ($p<0.0001$). There was no difference in the intercept nor frequency-dependent preference for yellow plants between observed and permuted datasets ($p\geq 0.2428$; figure 2.3A).

The probability of observing conspecific transition was significantly greater than expected under random foraging (figure 2.3B), reflected in the significant difference in intercepts between observed and permuted datasets ($p=0.0007$). The probability of conspecific transitions did not depend on the colour of the plants being visited ($p=0.4146$). The observed regression coefficient for the frequency-dependent component of assortative mating was shallower than in permuted datasets ($p=0.0056$). At the lowest frequency

variable	estimate	exp.	SE	ΔAIC	p-value
phenotype x frequency				-3	0.2245
frequency x density				-1	0.0756
phenotype x density				-5.1	0.4235
density				-1.5	0.4885
phenotype				+1	0.052
site				+1.1	0.08634
frequency	2.002	7.363	0.734	+4.9	0.0085
n. flowers	-0.009	0.991	0.003	+4.6	0.0058
n. tags	-0.237	0.788	0.046	+24.9	<0.0001
intercept	-1.183		0.653		

Table 2.1: GLMM analysis of tag drop data for wild plants. We removed variables explaining the least amount of variation stepwise in the order shown, and retained variables if this led to a significant increase in AIC. ΔAIC shows the change in AIC when a variable is removed along with the associated p-value. For retained variables, the table shows the estimate of the intercept, or of regression coefficients of the log odds ratio of observing a pollinator visit, along with the exponentiated value.

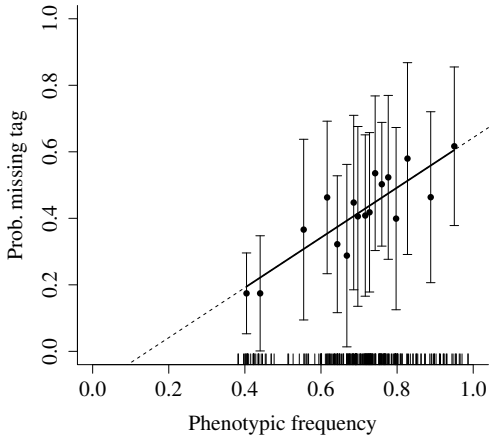


Figure 2.2: The probability of observing any tag to be missing on an individual per survey versus the phenotypic frequency of that plant. For clarity, data for individuals plants have been grouped into 15 bins of 17 individuals, with a mean with standard deviation for each group. Rugs show the distribution of individuals along the x-axis. To aid interpretation, a least-squares regression line has been fitted for binned values on phenotypic frequency. Where the line extrapolates beyond the range of observed data, this line is dashed.

	variable	observed	permuted	SE	p-value
A	intercept	-2.642	-2.376	$9.941 \text{x} 10^{-4}$	0.2428
	prev. yellow	0.601	-0.173	$9.996 \text{x} 10^{-4}$	<0.0001
	freq. yellow	4.485	4.888	$9.944 \text{x} 10^{-4}$	0.386
	var(bees)	0.528		0.6836	
	var(replicates)	$1.522 \text{ x} 10^{-10}$		0.0375	
B	intercept	-1.368	-2.514	0.227	0.0007
	prev. yellow	-0.236	-0.095	0.148	0.4146
	prev. frequency	3.900	4.988	0.373	0.0056
	var(bee)	0.286		0.535	
	var(replicate)	0.000		0.000	

Table 2.2: Analyses of pollinator behaviour in experimental arrays. Table shows regression parameter estimates for transitions to yellow plants (A), or for conspecific transitions (B). Prev. yellow and prev freq. are the colour and frequency of the previous plant visited. Freq. yellow is the frequency of yellow plants in the array. Var(bees) and var(replicate) refer to the variance components for individual foraging bouts and replicates. Estimates show intercepts, logistic regression coefficients, variance of random effects and standard errors for observed datasets, and appropriate means for parameters from permuted datasets. P-values show two-tailed probabilities that inferred parameters differ from parameters inferred for permuted datasets with flower colours randomly shuffled.

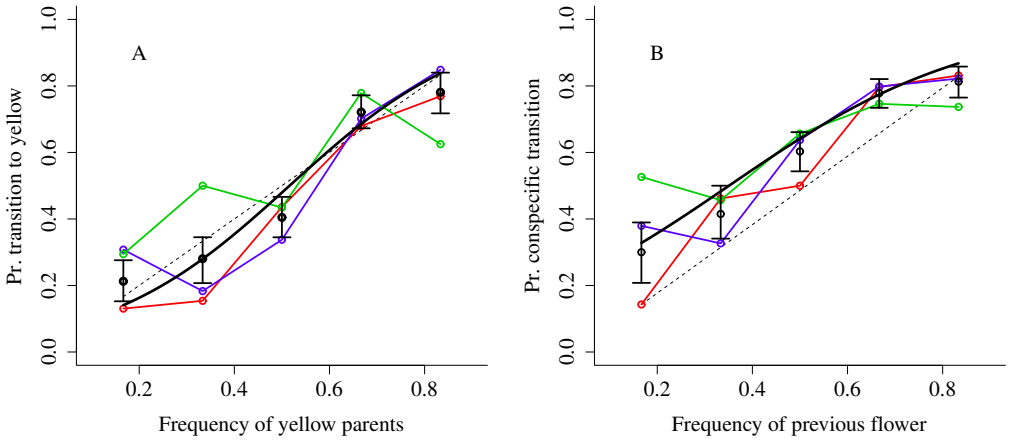


Figure 2.3: Pollinator transitions in experimental arrays, showing (A) pollinator preference for yellow flowers at different frequencies of yellow flowers, and (B) the probability that a pollinator transitions to a flower of the same colour as the plant she most recently visited. Black circles show the transitions across the whole dataset, and the thick solid line shows the curve predicted by the GLMM in table 2.2. Error bars show 95% confidence intervals derived from the means of 1000 bootstrap resamples. Red, green and blue lines show means for the August 2013, September 2013 and May 2014 replicates. Dashed lines show the expectation under random foraging.

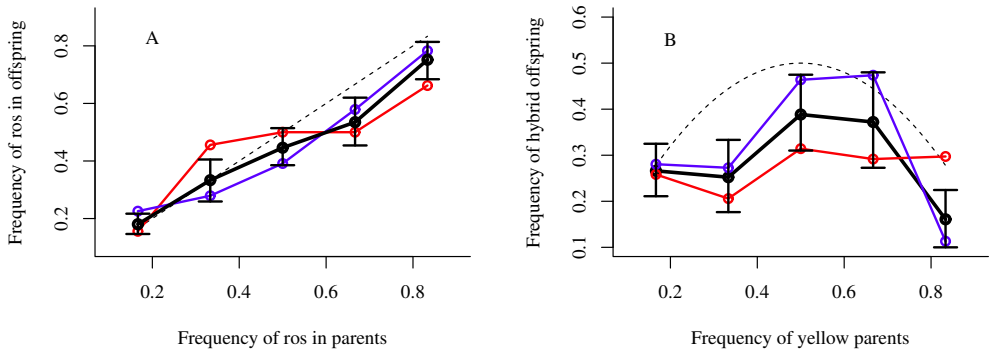


Figure 2.4: Allele and hybrid frequencies in progeny arrays. (A) Frequency of the non-pigmented *ros* allele in parent and progeny populations. (B) Frequency of hybrids in progeny populations. Red and blue lines show means for the August 2013 and May 2014 replicates, whilst the solid black line is the average across replicates. Error bars show 95% confidence intervals derived from the means of 1000 bootstrap resamples. Dashed lines show the expectation under random foraging.

of yellow plants the probability of a conspecific transition was 16.1% greater than under random mating, whilst when yellows were at their highest frequency this probability was only 3.5% greater than under random mating (figure 2.3B).

2.4.3 Deficit of hybrids in progeny arrays

We scored *ROSEA* genotypes of 712 progeny which survived to flower (151, 34, 70, 48 and 37 plants for the August 2013 replicate and 82, 77, 69, 38 and 108 plants for the May 2014 replicate for yellow:magenta frequency treatments of 6:30, 12:24, 18:18, 24:12 and 30:6 respectively). At low to medium frequencies of yellow parents, *ROSEA* allele frequencies in the progeny did not differ from expectations under Hardy-Weinberg equilibrium (figure 2.4A). At higher frequencies of yellow parents, there was a slight deficit of the non-pigmented allele. Overall observed allele frequencies do differ slightly from expectations under random mating ($\chi^2 = 4.45$, df=1, p=0.0349).

The average frequency of hybrid progeny was lower than expected under random mating for all treatments, except when yellow plants were very rare (figure 2.4B). This deviation from Hardy-Weinberg equilibrium was statistically significant (χ^2=23.95, df=1, p<0.001), although the deficit was generally much greater for the August 2013 replicate.

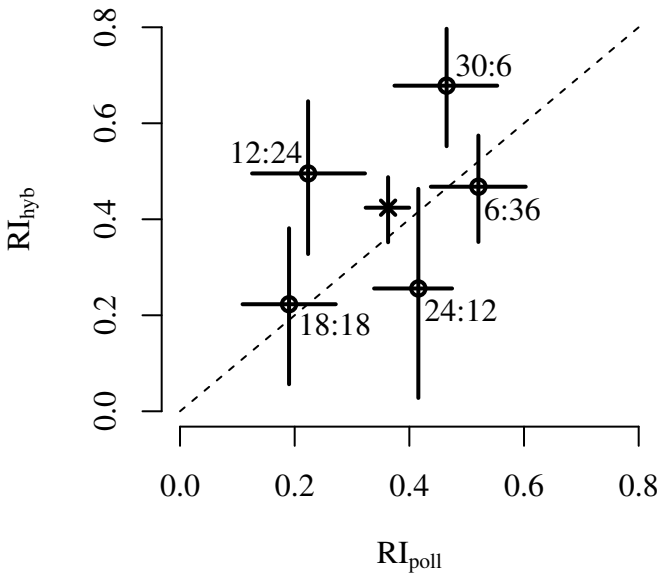


Figure 2.5: Reproductive isolation indices inferred from pollinator behaviour and hybrid frequencies in experimental arrays. The plot shows indices of reproductive isolation predicted by pollinator behaviour (RI_{poll}) and realised in the deficit of hybrid offspring (RI_{hyb}). Circles show means for each frequency treatment averaged across all three replicates for RI_{poll} and across August 2013 and May 2014 replicates for RI_{hyb}. Labels indicate the numbers of yellow to magenta plants in each treatment. The black cross is the global average across treatments. Error bars show 95% confidence intervals derived from the means of 1000 bootstrap resamples. The dashed line shows the expectation of 1:1 correspondence between RI_{poll} and RI_{hyb}.

There was considerable variation among arrays as to how well pollinator behaviour predicted assortative mating (figure 2.5). Across the whole experiment, mean isolation indices for pollinator behaviour and hybrid deficit were 0.363±0.020 and 0.424±0.035, indicating that pollinators were on average a good predictor of plant mating patterns in the arrays.

2.5 Discussion

This study has revealed several lines of evidence that suggest that bumblebees forage non-randomly on *Antirrhinum*. We used a tag-based assay of pollinator visitation to infer pollination success in the hybrid zone, which is robust to false positive errors. This assay revealed that the primary determinant of pollination success in this *Antirrhinum* hybrid zone is the composition of the floral neighbourhood. Over multiple seasons, plants were much more likely to be visited by a pollinator if they match the phenotype of as many neighbours as possible (figure 2.2). We found clear floral constancy by naïve bees in array experiments which was strong enough to explain a corresponding deficit in hybrid progeny in the offspring. These findings demonstrate that bumblebee pollinators forage non-randomly on *Antirrhinum,* and that this in turn causes positive frequency-dependent selection on and assortative mating among plants.

We found a discordance between the degree of frequency-dependent selection observed in the wild and in experimental arrays. Despite clear evidence for positive frequency-dependent visitation to wild plants (figure 2.2, table 2.1), we could not confirm this result in our array experiments (figure 2.3, table 2.2). Frequency-dependent foraging is highly sensitive to the density and spatial arrangement of flowers (Smithson and Macnair, 1997; Tastard et al., 2012). It is likely that we did not observe frequency-dependent behaviours in the arrays because the design of the experiment does not match the density or context of plants in the wild. The observation of constancy without concomitant frequency-dependence indicates that this constancy was caused by direct preference for familiar flowers, rather than indirect effects of frequency-dependence. Nevertheless, the increased visitation to locally common plants in the hybrid zone matches the correlation between frequency and seed set found by Tastard et al. (2012). This confirms that pollinators are the agent of this selection for common morphs, and that pollinators do indeed affect at least the female component of reproductive success. These

data do not allow us to evaluate the relative contribution to male fitness, but since this is not constrained by pollen limitation, increased pollinator visitation should be at least as important to male as to female fitness. Although previous studies have found increased fitness for common morphs in experimental arrays (Levin, 1972; Epperson and Clegg, 1987), this study and that of Tastard et al. (2012) are the first to demonstrate positive frequency dependent selection on wild plants, and show that pollinators are the causative agent.

Both pollinator observations and progeny frequencies showed that pollinators cause a marked drop in hybridisation between *A. m. striatum* and *A. m. pseudomajus* in our array experiments. Since we did not include hybrid plants in our arrays we cannot rule out the possibility that hybrids could act as a bridge to gene flow between the two subspecies. Nevertheless, this demonstrates that pollinators cause assortative mating based on flower colour in *A. majus.* It is difficult to account for the deficit in hybrids through F_1 inviability. Andalo et al. (2010) found no evidence of fitness differences between these subspecies and their F_1 hybrids, and we found that pollinator behaviour is sufficient to account for progeny phenotype frequencies (figure 2.5). Despite considerable variation among replicates, pollinator behaviour was a surprisingly good predictor of plant mating on average. Other array experiments have demonstrated that pollinators alone can cause almost complete reproductive isolation (e.g. Mather, 1947; Ramsey et al., 2003). However these studies compared full species which are divergent for many traits, so this is not surprising. Waser (1986) reported reproductive isolation indices for four comparisons of colour morphs within species and found values between 0.12 and 0.62. Applying this index to data on isolation between yellow and white morphs in *Raphanus raphanistrum* (Stanton, 1987a) gives an index of 0.22. Our estimates of reproductive isolation indices of 0.36-0.42 are within the range of these estimates. Interestingly, we found a negative correlation between constancy and the frequency of yellow plants. This could explain the finding by Tastard et al. (2012) that hybrid zone plants with strong magenta-pigmentation at high densities set more fruit when locally rare. A colour-specific frequency effect has also been found for bees foraging on *Ipomoea purpurea* and artificial flowers, where blue and pink flowers were visited even at low frequency, whilst white and yellow flowers were difficult to find when rare (Epperson and Clegg, 1987; Smithson and Macnair, 1997). The greater constancy observed in our experiments when yellows were rare may also be related to how well pollinators can pick out particular colours

from the floral neighbourhood at particular densities. These results highlight the importance of floral context in pollinator behaviour, and in turn on plant fitness.

The relationship between frequency and pollinator visitation we observed has clear implications for the dynamics of the *Antirrhinum majus* hybrid zone. It implies that pollinators learn to handle whichever morph is locally common, and preferentially visit that morph over rare phenotypes. On either side of the hybrid zone a single morph dominates at a frequency of $\sim$1, whilst a new migrant crossing into these areas will be at a frequency of $\sim$0. Our GLMM of visitation in the hybrid zone predicts that these locally common plants experience at least 37% greater reproductive success than the rare migrants (figure 2.2, table 2.1). The low frequency of migrants means they will be avoided by pollinators, and suffer reduced reproductive success, contributing substantially to the maintenance of the sharp clines at pigmentation loci. This mechanism is in contrast to that found in other plant hybrid zones, where separate species (or phyla) of pollinators show strong preferences for a suite of divergent floral traits (e.g. Campbell et al., 1997; Fulton and Hodges, 1999; Emms and Arnold, 2000; Streisfeld et al., 2013). Conversely, in this *Antirrhinum* hybrid zone pollinator service is performed by a single bumblebee species responding on a fine scale to changes in the floral community, rather than to differences between plants themselves.

Our array experiments suggest that we might expect magenta migrants into the yellow population to do better than would yellow migrants in the magenta population. Firstly, floral constancy was weaker when yellow plants were common, which makes it more likely that a pollinator would switch to a rare migrant (figure 2.3). Moreover, we found a slight increase in the pigmented *ROS* allele in array experiments when yellow plants were dominant relative to random mating (figure 2.4). In the hybrid zone, an advantage to rare magenta plants would facilitate introgression from *A. m. pseudomajus* into *A. m. striatum,* but restrict it in the opposite direction. Frequency-dependent behaviour by pollinators is therefore important both for maintaining the boundary between these two subspecies, and mediating patterns of gene flow between them.

The strength of selection inferred for wild plants is similarly strong to estimates of selection for other systems subject to positive frequency-dependent selection. *Heliconius* butterflies rely on warning colouration to deter predators, and pattern morphs form sharp hybrid zones. Using reciprocal transplant experiments, Mallet and Barton (1989b) showed that rare 'migrant' morphs

of *H. erato* had a 25-71% (mean=52%) drop in lifespan compared to the local morph. Similarly, in caged populations of *Partula suturalis* snails, the sinistral morph at a frequency of 10% produced 41% fewer offspring than the dextral morph at a frequency of 90% (Johnson, 1982). These three examples highlight the striking effect that positive frequency-dependent selection on a single trait can play in mediating reproductive isolation between populations.

Our findings highlight several open questions of broad relevance for our understanding of reproductive isolation. Firstly, more detailed analysis of the clines in allele frequencies across the hybrid zone would allow us to relate fine-scale selection on individuals to selection over a longer term. Secondly, a better understanding is needed of the mating patterns between individuals in the hybrid zone. This would shed light on the relative contributions of frequency-dependent selection and assortative mating to cline maintenance, and in turn on the behavioural patterns of the pollinators which cause it. Furthermore, reconstructing mating patterns would elucidate the relative contribution of pollinator behaviour to male and female components of reproductive success. These questions are of broad relevance to our understanding of reproductive isolation between plant populations in general.

2.6 Collaborator contributions

Annual surveying of the hybrid zone was organised by Enrico Coen's group and David Field, supported by a large number of volunteers. I participated in 2011, 2012, 2014 and 2015. Enrico Coen's group designed the tag assay experiment, and collected data in 2010; subsequent data collection was performed by myself and nine volunteers. Seeds for experimental arrays were collected and crossed by Coline Jaworski (Jaworski et al., 2015).

Chapter 3

Joint estimation of paternity and sibship structure for half sibling families

3.1 Abstract

Knowledge of genealogical structure is of great utility for understanding the genetics of wild populations. Considering full-sibling families as a whole increases statistical power to assign parentage, but this requires that these full sibships can be identified. There are typically an enormous number of ways to partition a set of individuals into family groups, so heuristic methods are needed to sample these partitions efficiently. However, relatedness between candidate parents means that genetic data are often consistent with multiple possible genealogies, and small reconfigurations to these family groups can lead to large changes in likelihood. We describe an efficient method to estimate the likelihood of full sibships considering all fathers simultaneously. Secondly, we incorporate an informative prior distribution on family size, which increases the accuracy of sibship assignment. We also describe a fast and accurate clustering method to explore possible family structures. By testing these methods against simulated datasets we show that our method increases the accuracy and speed of sibship inference relative to existing methods.

3.2 Introduction

The genealogical pedigree of a population is a valuable piece of information for genetic studies because it conveys complete information about the relatedness between individuals, as well as the mating patterns and variance in reproductive success within a population (Pemberton, 2008). For humans and domesticated organisms this can often be reconstructed directly from civic or breeders records, but for wild populations this typically has to be inferred from from marker data (Blouin, 2003; Jones et al., 2010). Since the entire pedigree is challenging to estimate directly, this has typically involved the estimation of the most likely pairwise relationships between individuals under Mendelian segregation, which can then be combined into a single genealogy (Thompson, 1975, 1976b). A major focus of such work has been the estimation of parent-offspring relationships, which seek to find one or both parents from a pool of candidates (Meagher, 1986; Meagher and Thompson, 1986; Devlin et al., 1988; Marshall et al., 1998; Hadfield et al., 2006). A second group of methods have developed which aim to find the best partition of individuals into full sibship groups (FSG) when information about the specific parents is not available (Painter, 1997; Thomas and Hill, 2000; Emery et al., 2001; Smith et al., 2001; Thomas and Hill, 2002; Wang, 2004). However, considerable gains in statistical power can be gained by considering family groups as a whole (Wang, 2007; Wang and Santure, 2009) because including more individuals provides extra information on the complete set of alleles in the family. Thus, combined approaches which jointly estimate FSGs and their parentage are powerful tools for piecing together pedigrees in the wild.

Methods to assign sibship structure can be differentiated by two aspects of their algorithm. Firstly, some kind of score is needed to compare the degree of support for different hypothesised partitions. Secondly, because the space of possible partitions is too large to be enumerated exhaustively, some kind of heuristic is needed to explore this space efficiently. Most approaches achieve this by multiplying over the likelihood of observing each individual in a putative FSG given all possible parental allele contributions (Painter, 1997; Thomas and Hill, 2000; Emery et al., 2001; Smith et al., 2001; Wang, 2004), or observed parental genotypes (Wang and Santure, 2009), and then multiplying over the likelihood for each sibship. They then use an iterative algorithm to ascend the likelihood surface to identify the best-fitting partition(s), such as simulated annealing (Wang, 2004; Wang and Santure, 2009;

Thomas and Hill, 2000) or a Monte Carlo Markov Chain (MCMC; Emery et al., 2001; Thomas and Hill, 2002; Smith et al., 2001). We will henceforth refer to such algorithms as 'hill-climbing' algorithms, although we note that they can deal with some degree of unevenness in the likelihood surface. The most sophisticated of these programs, Colony, sequentially reassigns parents and partition structures as it ascends the likelihood surface, meaning that parents are considered in turn (Wang and Santure, 2009; Wang, 2012). This means that when there are many candidate fathers it may take a very long time for the true father to be considered, and furthermore the process of comparing support for different parentage assignments is intimately linked to the process of exploring partition space. One notable exception is the approach of Almudevar and Field (1999), which first identifies a set of plausible FSGs using the principle of exclusion of opposing homozygous genotypes, and then compares these using a scoring system based on segregation ratios. Thus, aside from the latter method, current approaches rely on some kind of hill-climbing procedure to explore partition space, and only Wang and Santure (2009) incorporate parental genotype information into FSG inference.

Genealogy inference is complicated by relatedness among candidate parents. For example the uncle of a FSG will share alleles with the true father because they are brothers. Due to stochasticity in Mendelian sampling may have a higher likelihood of paternity of some of his nephews than does the true father, a paradox noted by Thompson (1976a). Genetic data can therefore be consistent with one or more incorrect genealogies, which in this case causes FSGs to be split into multiple subsets. There is growing evidence that the small changes to partitions can cause large differences in support for different models (Butler et al., 2004; Wang and Santure, 2009; Almudevar and Anderson, 2012). This has two important implications for the two algorithm components described above. Firstly, this will introduce local maxima into the likelihood surface. Iterative approaches depend on smooth likelihood surfaces to climb, so dramatic variation in likelihood surface topology may make it difficult to locate the true partition. Although simulated annealing and MCMC are designed to escape local maxima, they are inefficient when differences in likelihood are large. Secondly, support for the true partition can be swamped by support for incorrect partitions. This will mean that even in the area of the global maximum, algorithms can bypass the correct partition and arrive at a related, but incorrect, partition. The extent to which is a problem will depend on the extent to which alternative partitions structures differ from the true structure and how this relates to the biological question.

Simulation studies can be used to test the performance of algorithms when the true partition is known. Such studies have shown that algorithms frequently split up real FSGs and/or spuriously assign unrelated individuals to the same FSG (Smith et al., 2001; Butler et al., 2004; Wang, 2004; Wang and Santure, 2009; Almudevar and Anderson, 2012), although the biases were different for different programs and different underlying data structures. The spurious inferences are dramatically exacerbated when genotyping and human errors were introduced into the dataset (Almudevar and Anderson, 2012). Two developments are therefore needed to overcome the problems associated with rough likelihood surfaces in sibship inference: (1) a measure which gives the greatest support to the true genealogical structure, and (2) a way of exploring the space of likely structures which does not depend on an iterative algorithm.

In this paper we develop a new method for jointly estimating FSGs and their paternity for arrays of full and half siblings. Our approach uses information about the paternity of individuals to infer a joint distribution of the paternity for whole FSGs. We describe a Monte Carlo simulation approach to estimate the likelihood of proposed partitions quickly and accurately, even for very large samples of males. We propose the use of prior probabilities of particular partitions to weight inference towards larger or smaller sibship sizes to try and alleviate the effect of greater support for incorrect genealogies. We then describe a similarity-based clustering algorithm to explore the space of possible partitions which does not rely on any hill climbing algorithm and finally we test the performance of the method against simulation.

3.3 Methods

An array of half siblings consists of a set of offspring individuals $O = o_1, o_2, \ldots, o_n$ sharing a single mother m. Assuming all fathers have been sampled, each offspring individual has a single father from among the set of candidate fathers $F = f_1, f_2, \ldots, f_n$. Inference of sibship structure relies on matrix $\mathbf{G}$, where each row of $\mathbf{G}$ corresponds to a single offspring individual and each column to a candidate father. Element g_{ik} of $\mathbf{G}$ represents the posterior probability that the k^{th} candidate father in F is the true father of the i^{th} offspring in O. Because each row represents a probability distribution, the rows of $\mathbf{G}$ must sum to one (Devlin et al., 1988; Nielsen et al., 2001), and must also include the probability that a real father has not been sampled (see

Appendix). Methods to assign a likelihood of paternity based on shared alleles are well established (Thompson, 1976b; Meagher and Thompson, 1986; Meagher, 1986; Devlin et al., 1988; Marshall et al., 1998; Nielsen et al., 2001; Anderson and Garza, 2006; Hadfield et al., 2006; see Jones et al. (2010) for review). We will therefore take estimation of $\mathbf{G}$ as given, except that we set the probability of particularly unlikely candidate fathers to zero a priori. We identify such excluded candidates as any males with more than a threshold number of incompatible homozygous loci with a given offspring. By keeping only 'compatible' candidates we remove background noise from $\mathbf{G}$ which greatly simplifies the task of likelihood estimation via Monte Carlo simulation described in the next section. This introduces the danger that the real father might be excluded from the dataset. This threshold should therefore be set appropriately given the total number of loci typed and expected genotyping error rates such that this occurs only rarely. This risk is more than outweighed by the boost in power gained by simplifying $\mathbf{G}$.

Sibship assignment can be thought of as finding the best partition C which groups the set of individuals in a half sibling array O into subsets of valid FSGs. Set $x \subseteq O$ is a valid FSG if and only if all individuals in x share the same father. The posterior probability of C depends on the likelihood of observing the maternal, paternal and offspring genotypes given C, and the on the prior belief in that C:

$$\Pr(C|\mathbf{G}) \propto \Pr(\mathbf{G}|C)\Pr(C). \tag{3.1}$$

In the next two sections we describe an efficient method for estimating $\Pr(\mathbf{G}|C)$ by using $\mathbf{G}$ to construct a joint distribution of paternity for full sibships, and suggest an appropriate form for the prior distribution $\Pr(C)$ on sibship size.

3.3.1 Likelihood of a partition

Let γ denote the matrix of probabilities of paternity for each sibship, which like $\mathbf{G}$ has one column for every candidate father, but whose rows correspond to proposed full sibships rather than individuals. For the putative FSG x given partition C, the likelihood that each father is the true father of every individual in x is the product of his probability of paternity for each offspring. We can therefore define γ_x, the vector of paternity likelihoods for x as the element-wise product of each row in $\mathbf{G}$ for all individuals in x. Since the

true father of x may be any one of the candidate fathers, the probability of x is proportional to the sum over γ_x. Assuming each sibship is independent the likelihood of the proposed partition is then simply the product of the likelihoods for each set of full siblings:

$$\Pr(\mathbf{G}|C) = \prod_i \sum_k \gamma_{ik} \tag{3.2}$$

where i *i*ndexes each full sibship in *X, k* indexes candidate fathers.

However, because two full sibships cannot by definition share the same father, this formulation is only valid if the set of compatible fathers for each full sibship is unique, with no intersection with that of any other full sibship. To visualise this we can view equation 3.2 as a product of sums, and hence the whole expression as an algebraic expansion. For example, imagine two proposed full sibships, *a* and *b,* and three candidate fathers, such that $\gamma_a = \{a_1, a_2, a_3\}$ and $\gamma_b = \{b_1, b_2, b_3\}$, where the subscript indicates the likelihood for each candidate. Equation 3.2 is then $\sum_k \gamma_{ak} \sum_k \gamma_{bk} = (a_1 + a_2 + a_3)(b_1 + b_2 + b_3)$, which expands to a sum of nine terms for all possible pairwise combinations:

$$\begin{array}{ccc} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{array}$$

To account for terms describing shared fathers across sibships it is necessary to remove those terms with one or more matching subscripts, which in this example are the diagonal elements. For small cases such as this it is straightforward to identify these terms and subtract them from equation 3.2, but this quickly becomes computationally demanding as the number of sibships and candidate fathers increases (with five sibships and 100 candidate fathers there are $7.88\text{x}10^{69}$ terms to enumerate, for example).

Nevertheless, $\Pr(\mathbf{G}|C)$ can be efficiently estimated using Monte Carlo simulations. Because each sibship has exactly one father, we can sample a valid set of possible fathers for each sibship by traversing a path through γ from top to bottom, visiting each column at most once, and each row exactly once (figure 3.1). The likelihood of each path through γ is the product of the probability of each element visited, and the likelihood for the whole partition is the sum of likelihoods for all valid paths through γ. As before there will be generally be too many paths to enumerate. However, because

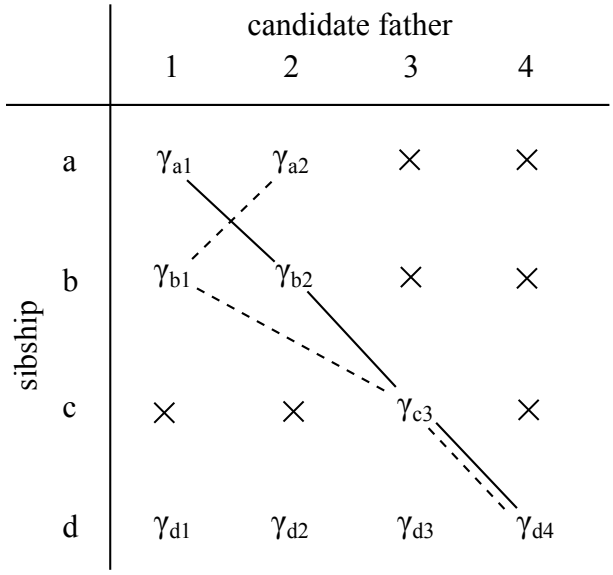


Figure 3.1: Valid paths through matrix γ. An example of matrix γ for four sibships and four possible candidate fathers compatible with at least one of the sibships. Crosses indicate males excluded based on opposing homozygous loci, whose probability has been set to zero. A set of fathers is drawn by traversing the graph from top to bottom, visiting each column no more than once. The solid and dashed lines represent the only two valid paths through the matrix ($\{\gamma_{a1}, \gamma_{b2}, \gamma_{c3}, \gamma_{d4}\}$ and $\{\gamma_{a2}, \gamma_{b1}, \gamma_{c3}, \gamma_{d4}\}$ respectively). Note that paths through the graph are much simpler to enumerate once unlikely fathers have been removed.

most rows will contain one or a handful of large values and many small values, only a small number of the possible paths will explain most of the total probability, and we can approximate the likelihood of the partition by only considering these. To do this we draw samples with replacement from each row of γ, sampling proportionally to the probability of each element. Assuming each sibship is independent, we can combine drawn fathers into sets representing paths through γ. We then remove duplicate sets, and any sets where two or more sibships share a father. The paths that remain are sets of unique fathers representing unique paths through γ, drawn proportionally to the likelihood of those paths. We have found sampling to be very fast; 10,000 Monte Carlo draws for 25 sibships with 1000 candidate fathers took 5.29ms on a MacBook Pro with a 2.4 GHz Intel Core i5 processor. It should be noted that this approximation may underestimate the true likelihood if many possible paths are possible, but given sufficient Monte Carlo draws this underestimate will be much smaller than differences between the likelihoods for different proposed partitions, and should therefore have a negligible effect on the power to distinguish these.

3.3.2 Prior probability of a partition

Ideally we wish to make as few prior assumptions as possible regarding the structure of partitions and let inference be guided by the signal from the data. Previous investigations of sibship inference have, explicitly or implicitly, used flat priors to weight each structure equally and rely only on the likelihood (Thomas and Hill, 2000,2002; Wang 2004; Wang and Santure 2009). It has frequently been observed that sibship inference algorithms tend to split sibships into smaller FSGs, even where the true structure is known, or there are biological reasons to expect larger sibship sizes (Thomas and Hill, 2000; Smith et al., 2001; Wang, 2004; Butler et al., 2004; Almudevar and Anderson, 2012). This suggests that relying on the likelihood alone exacerbates the effect of Thompson's paradox (Thompson, 1976a). Informative prior distributions can be used to weight the posterior probability of particular hypotheses when we have sound prior reasons to do so. We propose the use of the geometric distribution to weight the probability of smaller or larger sibships. The geometric distribution describes the probability of observing a successful Bernoulli trial after $n-1$ failures, where the probability of success in each trial is p. If n_i describes the number of individuals in full sibship x_i the prior probability of the i^{th} full sibship is

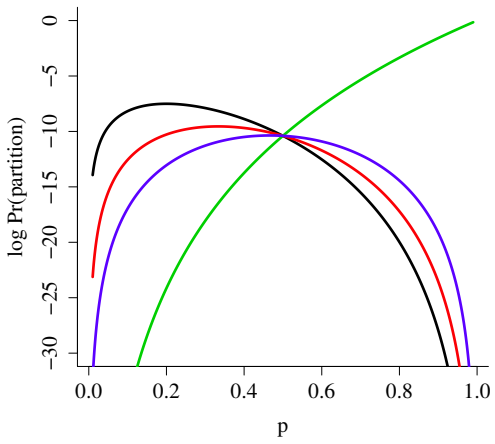


Figure 3.2: Examples of log prior probabilities of four partition structures under a geometric prior distribution given p between zero and one. Lines indicate three FSGs of five siblings (black), five FSGs of three siblings (red), 15 FSGs of 1 siblings (green). The blue line indicates a skewed family distribution made up of two FSGs of five siblings plus five FSGs of one siblings.

$$\Pr(x_i) = (1-p)^{1-n_i}p. \tag{3.3}$$

The prior probability of the partition is then the product over each full sibship, multiplied by any prior on the total number of sibships $\Pr(N)$ if appropriate:

$$\Pr(C) = \Pr(N) \prod_i \Pr(x_i). \tag{3.4}$$

Low values of p will give greater weight to fewer, larger sibships and *vice versa* (figure 3.2).

3.3.3 Exploring partition space

Agglomerative clustering algorithms are a very fast approach to clustering items into groups. These algorithms apply a 'linkage function' to a matrix of distance (or similarity) measures to identify the most similar clusters.

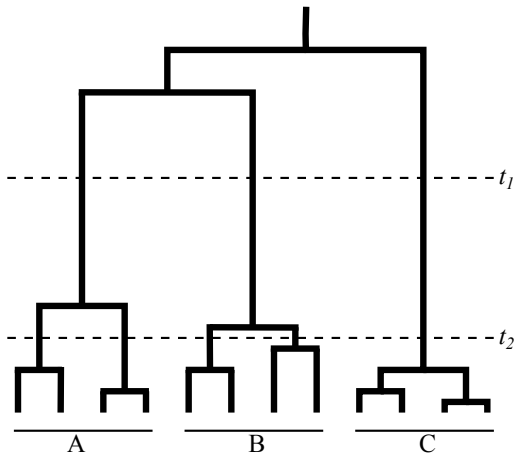


Figure 3.3: Example of a bifurcating dendrogram of relatedness between twelve individuals in three full sibling families. The actual number of clusters depends the horizontal threshold at which branches are declared to represent true clusters. For example, at if the threshold is set at t_1 there would be three clusters, but at t_2 there would be five clusters.

The Unweighted Pair Group Method with Arithmetic Mean Algorithm (UPGMA) (Sokal and Michener, 1958), for example, takes the average distance between all pairs of elements in each cluster. These algorithms begin with every individual in a cluster on its own, and then sequentially join the most similar pair of clusters until all individuals have been grouped. This results in a bifurcating dendrogram (figure 3.3), with closely related individuals sharing branching close to the tips of the tree, and less closely related individuals branching towards the root. Taking slices through the dendrogram at different distance thresholds will give different clusters. Because these algorithms depend only on the distance matrix, they provide a slice through the partition space without having to ascend a probability surface.

For the case of sibship inference, we estimate a distance matrix $\mathbf{D}$ whose ij^{th} element is based on the likelihood $\sum_k g_{ik}g_{ik}$ that the i^{th} and j^{th} offspring are full siblings, where k indexes each father in F. There are a number of ways to convert this into a distance matrix, but we use the absolute values of the log likelihoods:

$$\mathbf{D_{ij}} = \left| \ln \sum_k g_{ik}g_{ik} \right|. \tag{3.5}$$

We then use the UPGMA (Sokal and Michener, 1958) to build a dendrogram of individuals based on **D.** It is then straightforward to assess the posterior probability of the clusters implied at all possible thresholds on the clustering dendrogram (figure 3.3).

Nevertheless, this clustering approach is simply a slice through partition space which may not intersect the global optimum. As noted above, hill climbing algorithms can be used to ascend likelihood surfaces, but get stuck in local optima. A compromise position is to use clustering to reach an area of high posterior probability, and use hill-climbing to 'fine-tune' the position in partition space from there. Starting from partition C, we consider all new partitions which can be made by moving a single individual into every other available FSG, including into a FSG on its own. We assess the posterior probability of each, and accept the partition with the greatest support. This is repeated until no new partitions can be made with greater support than the current partition.

3.3.4 Simulations

To assess the performance of the methods we used the methods described above to infer FSGs in half-sibling arrays of simulated offspring. We simulated 200 replicates of 1000 diploid candidate parents with 100 SNP loci using minor allele frequency distributions drawn from a uniform distribution between 0.3 and 0.5. For each half-sibling array we drew three families of five offspring following Mendelian inheritance from a single mother and three independent fathers. To mimic real world datasets, we also introduced genotyping errors at a rate of $\epsilon_1 = 0.01$ per haploid nucleotide and dropouts at $\epsilon_2 = 0.01$ per diploid locus. We estimated $\mathbf{G}$ given the marker data and error rates following Anderson and Garza (2006), but account for missing SNP data by taking the geometric mean over loci (Appendix). We did not remove any true fathers from the dataset, and set the proportion of missing fathers to zero when estimating $\mathbf{G}$. Simulations were carried out using scripts written in Python.

For each dataset we used the UPGMA to identify the set of possible partitions based on distance matrix D, and recorded how often the most probable partition matched the true partition. We then used these most probable partitions as the starting point for hill-climbing by steepest ascent to identify any nearby optima not captured by the clustering algorithm. Initial simulations using hill-climbing alone very quickly became stuck in

local maxima far from the true partition, so we did not pursue this approach in subsequent simulations. We estimated the likelihood of partitions using 10,000 Monte Carlo draws as described above, and using either uninformative (flat) priors or a geometric prior with p=0.2 (figure 3.2). In all cases, we also set $\Pr(N)$ to 0 for $N = 1$ and 1 for $N \geq 2$ to ensure arrays were separated into at least two families. As a measure of accuracy we used the proportion of the 105 pairwise relationships among the 15 individuals which were correctly identified as a full- or half-sibling relationship. We also quantified the sensitivity to two kinds of underlying errors: (1) the frequency with which a single individual splinters from the true FSG into a family on its own, and (2) the frequency with which two real FSGs are merged into a single FSG. To assess the performance of the method given different numbers of loci, we performed the calculation of $\mathbf{G}$ and subsequent sibship inference for the full 100-locus dataset, as well as subsets of each dataset with 40, 60 and 80 loci.

Finally we used the simulated datasets to assess the difference in support for the true and most likely partitions. We did this by comparing the log posterior probability of the true partition ($\Pr(C|\mathbf{G})_{true}$) to the best-supported partition that could be found by clustering and hill-climbing ($\Pr(C|\mathbf{G})_{max}$), although we note that this may itself represent a local rather than global optimum.

3.4 Results

3.4.1 Accuracy of partitions

For every dataset examined the true partition was among the possible set of partitions inferred, although this was not necessarily the partition with the highest posterior probability. The accuracy of simulations depended surprisingly little on the number of loci in the dataset. There is no clear trend in accuracy as the number of loci increases in figure 3.4, although the variance for 40 loci when using geometric priors is substantially larger than for larger datasets.

Using flat priors, a median of 88.6% of pairwise relationships were inferred correctly using both clustering alone and clustering plus hill-climbing (figure 3.4). Including hill climbing led to an increase in accuracy in only 18.1% of datasets. Using flat priors at least one 'splintered' family of size one was

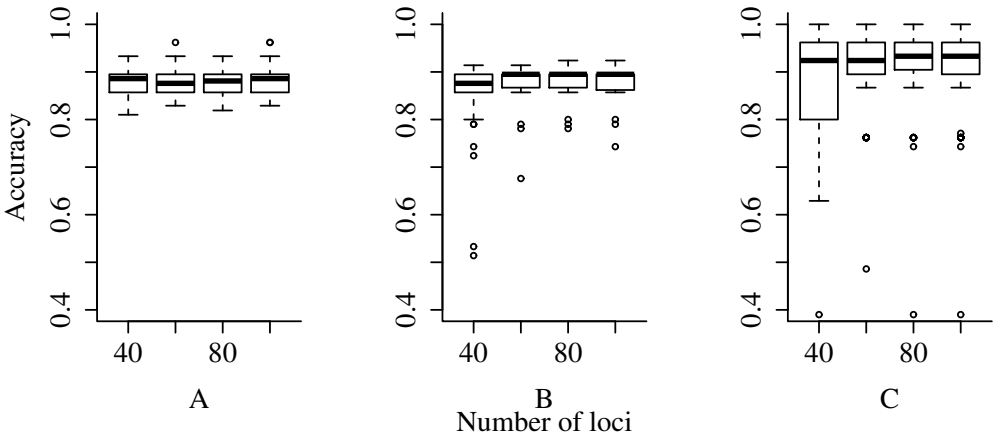


Figure 3.4: Accuracy of sibship inference for simulated datasets. Box plots show the proportion of correctly inferred pairwise relationships for using clustering alone with flat priors (A), clustering followed by hill-climbing with flat priors (B) and clustering alone with geometric priors (C). Distributions for clustering followed by hill-climbing with geometric priors were identical to C.

observed in every data set (mode=3 singleton families), but no instances of merging were observed.

Using geometric priors median accuracy increased to 93.3%, although the variance also increased (figure 3.4C). The correct partition was identified in 14% of datasets. Nevertheless, using geometric priors led to an increase in accuracy in 70.5% of datasets relative to flat priors. Hill-climbing never located any better-fitting partition, and accuracies before and after hill-climbing were identical. Using geometric priors reduced the occurrence of splitting to 72% of inferred partitions. However, it also increases the rate of merging two or three true FSGs into a single FSG from zero to 12% of datasets, which accounts for the increased variance in accuracies. Merging two or three FSGs account for the clusters of points at ~0.75 and 0.4 respectively in figure 3.4C.

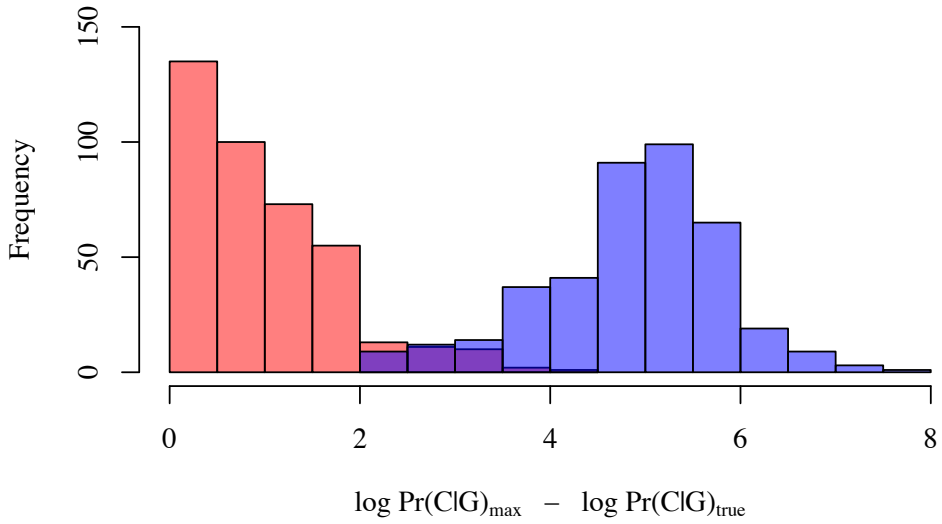


Figure 3.5: Differences in log posterior probability of the true partition and the most probable partition for 200 simulated datasets of three FSGs of five offspring each. Plot compares partition structures found via clustering followed by hill climbing using flat (red) or geometric (blue) priors.

3.4.2 Differences in support

In all datasets the algorithm found partitions with higher posterior probability than the true partition. Support for the most probable partition was on average 4.9 log probability units higher than for the true partition using flat priors, and 0.96 units higher using geometric priors (figure 3.5). In linear space this means that the true partition was on average 2.62- and 130-fold less probable than the best solution respectively.

3.5 Discussion

Despite being first noted almost forty years ago (Thompson, 1976a), the degree of unpredictability in the topology of likelihood surfaces has received

remarkably little attention in studies of genealogical inference. This study is the first to quantify the severity of this problem for the case of sibship and paternity inference, and to develop methods which address it. Our simulations show that we should in general expect the most likely partition to be incorrect, and furthermore that the relative support for the incorrect partition will be large enough to effectively exclude the correct partition from a sample of possible partitions most of the time. Ascending the probability surface by steepest ascent identified partitions with higher posterior probability, but provided no additional accuracy, showing that the surface topology exhibits many local maxima which will confound iterative algorithms. Nevertheless, this study demonstrates that the use of an appropriate prior distribution can be used to weight the probability surface in favour of larger family sizes, and improve accuracy of sibship inference. Furthermore, the use of a clustering algorithm to identify partitions removes the need to explore partition space iteratively, thus removing the dependence on local surface topology. Together, these methods lead to an improvement in sibship inference in more than 70% of cases, even in the presence of genotyping errors.

At the core of the methods described in this study is the matrix of pairwise probabilities of paternity $\mathbf{G}$. The precision of genealogical inferences are subject to noise caused by stochasticity in Mendelian inheritance, and an unrelated male can have a higher likelihood than the true father (Thompson, 1976a). Likelihoods of pairwise parent-offspring relationships are especially sensitive to this because only half the parental contribution is observed in the offspring (Wang, 2007). Nevertheless, we found that the true partition was among the possible partitions that could be inferred from the clustering dendrogram in every dataset examined, without the need for time-consuming iterative algorithms. This demonstrates that $\mathbf{G}$, and hence also relationship matrix $\mathbf{D}$, contain sufficient information to identify a path through partition space which includes the correct partition, at least for these well-behaved simulated datasets. How well this power is reflected in real world datasets will depend on the degree of linkage between markers, and on the structure of relatedness in the pool of candidate fathers. Nevertheless, our simulations showed that the observed errors in partition inference were not due to how UPGMA explores partition space, but rather to the measure of support used to compare possible partitions.

Our estimator for the likelihood of a partition has a number of advantages over other methods. Firstly it explicitly incorporates information on both parental and offspring genotypes into family reconstruction, and can deal

with genotyping errors (Butler et al. (2004); c.f. Painter 1997; Almudevar and Field 1999; Thomas and Hill 2000; 2002; Smith et al. (2001)). Considering all males simultaneously rather than sampling them iteratively decouples the process of paternity estimation from partition exploration and eliminates the possibility that the true father is never sampled by the algorithm (Wang and Santure, 2009). Finally, because we can never expect to estimate the true genealogy with complete accuracy, downstream analyses need to account for uncertainty in pedigree structure. Our method outputs a vector of probabilities that each father is the true father of a sibship, as well as a posterior probability for each partition, which means that this uncertainty can be integrated out. However, although we found no evidence that the likelihood estimator caused unrelated individuals to be inferred as full siblings, it did frequently cause individuals to splinter into a FSG of a singel individual. It seems that the propensity for splitting is linked to the number of candidate males which could not be excluded a priori from **G**. When there are many compatible males and the differences in posterior probability are not large, sum of the posterior probabilities for all incorrect males will often be greater than that of the true father. In this case the Monte Carlo sampler will draw the true father only rarely, even when geometric priors are used. This again highlights the sensitivity of pairwise estimators to Thompson's paradox. One approach to deal with this would be to set a very low prior probabilities for singleton FSGs, but this could cause merging if some true fathers are missing from the data set. A post-hoc approach would be to reassign singletons to large FSGs based on the support for shared compatible fathers. However, probably the most robust way to deal with splintering is to maximise the amount of information in **G** to restrict the number of compatible fathers as far as possible.

Our simulations show that using an informative prior distribution can make the probability surface more tractable and improve partition inference. Although it does not completely overcome the problem of sibship splintering, using a geometric prior distribution for sibship size reduced the differences in support for the true and most probable partitions by around 50-fold on average relative to using flat priors. This was accompanied by an overall increase in accuracy, and in some cases perfect reconstruction of sibship topology. Our results show that the geometric distribution is suitable when several large families of equal size are present, but other prior distributions may be more appropriate if small or highly-skewed FSG sizes are expected (Butler et al., 2004). The value of p of 0.2 used in our simulations represents a rather

strong prior, but this is necessary to deal with the large differences between $\Pr(C|\mathbf{G})_{true}$ and $\Pr(C|\mathbf{G})_{max}$ (figure 3.5). Nevertheless we frequently observed that weighting large sibships caused merging of two or more FSGs, showing that there is a fine balance between setting a sufficient prior on sibship size to limit splintering, and setting a sufficiently relaxed prior on the number of FSGs to prevent merging.

Despite issues with splintering and merging, our results compare favourably with those estimated for other methods. In a comparison of four sibship inference algorithms, Butler (2004) reports that genotyping errors caused 24-54% of previously correctly assigned individuals to be assigned into singleton FSGs (but see also the comment on this paper byAlmudevar and Anderson, 2012). Individuals were assigned to an incorrect sibship 26-72% of the time, whilst our method caused merging only 12% of the time using geometric priors and never when using flat priors. It should be noted that the simulation conditions used in that study were different from those used here, and that the authors found that the underlying sibship structure had a considerable effect on the performance of each algorithm. Furthermore, our Monte Carlo simulator and clustering method are extremely fast, and avoid the need to time-consuming iterative algorithms. The methods described here therefore represent an improvement over previous sibship inference methods both in terms of speed and accuracy.

The results of this study suggest several useful lines of further investigation and potential areas of improvement. Further investigation is needed to test the performance of the methods to different partition structures, and investigate sensitivity to misspecified prior distributions and missing fathers. The speed of clustering means that is feasible to repeat it many times with different hypotheses, for example regarding the proportion of missing fathers, or incorporating phenotypic or geographic information (Hadfield et al., 2006). At present the methods described here refer to half-sibling arrays only and assume that the mother is known with certainty. However it would in principle be straightforward to extend the method to full parentage assignment by modifying $\mathbf{G}$ to apply to putative pairs of parents rather than individual fathers. The application of the methods described in this paper and their possible extensions provide rich opportunities for improving our understanding of pedigree structures in the wild.

3.6 Appendix

3.6.1 Likelihood of paternity with SNP data

The likelihood that a male is the true father of an offspring is given by the probability λ_l of observing the offspring genotype from the maternal and paternal alleles at locus l, multiplied across each locus, so that $g_{ik} \propto \prod_l \lambda_l$ (Meagher, 1986; Meagher and Thompson, 1986). Anderson and Garza (2006) modified these methods to deal efficiently with SNP data. This includes dealing with point mutations, where a haploid genotype is observed to be allele A when it is actually allele B, by summing over all possible maternal, paternal and offspring alleles, weighted by the probability that each is the true genotype given an estimate of the error rate ϵ_1 , but failing to account for SNP dropouts causes candidate males with many failed loci to have high likelihoods of paternity, because the calculation multiplies over fewer loci. We account for these errors by correcting for the v loci which amplified successfully for a given maternal, paternal and offspring genotype, giving $g_{ik} \propto \prod_l \lambda_l^{1/v}$, or equivalently $\log g_{ik} \propto {}^1/_v \sum_l \log \lambda_l$. A second kind of genotyping error occurs when data for individual SNPs fail to amplify and are missing entirely. This drop-out rate, ϵ_2, can be observed directly from the data.

3.6.2 Missing fathers

The Monte Carlo simulation approach to estimating $\Pr(C|\mathbf{G})$, requires that each row of $\mathbf{G}$ sum to one. If all fathers have been sampled, each element of the i^{th} row if $\mathbf{G}$ is simply

$$g_{ik} = \frac{\Pr(o_i|f_k)}{\sum_k \Pr(o_i|f_k)}.$$

In real datasets it is unlikely that every male can be sampled, and therefore some offspring will have paternal genotypes not found in F. It is then necessary to modify $\mathbf{G}$ to account for the prior probability $\theta_i = \Pr(f_i \in \mathbf{f})$ that the father of the i^{th} offspring has been sampled and $\Pr(o_i|m, \mathbf{a})$ that an individual is the offspring of an unsampled father with alleles drawn at random from vector of local allele frequencies, $\mathbf{a}$ (Nielsen et al., 2001). Thus we have

$$g_{ik} = \frac{\theta_i \Pr(o_i|f_k)}{\sum_k \theta_i \Pr(o_i|f_k) + (1-\theta_i) \Pr(o_i|\mathbf{a})}.$$

To allow Monte Carlo simulations to sample configurations where one or more FSGs has a missing father we also append each row in $\mathbf{G}$ with

$$g_{ia} = \frac{(1-\theta_i)\Pr(o_i|\mathbf{a})}{\sum_k \theta_i \Pr(o_i|f_k) + (1-\theta_i)\Pr(o_i|\mathbf{a})}.$$

Individuals with large terms for g_{ia} may either be full siblings sharing a single unsampled father or else half siblings with multiple unsampled fathers, but it is difficult to distinguish these hypotheses in the absence for further information.

Chapter 4

Direct measures of mating patterns and reproductive success in a snapdragon hybrid zone

4.1 Abstract

The reconstruction of mating events is a valuable tool in evolutionary ecology because it permits direct estimation of variance in reproductive success and patterns of mating among individuals. The snapdragons *Antirrhinum majus pseudomajus* and *A. m. striatum* form a narrow hybrid zone in the Pyrenees. Here, the sharp transition in flower colour suggests strong selection on loci controlling magenta and yellow floral pigmentation. Previous work found that pollinators prefer locally common flower colours, and show floral constancy to particular colours. In this study we used paternity analysis of open-pollinated seeds to assess the direct impact of pollinator behaviour on mating among wild plants. Our approach accounts for uncertainty in paternal identity and relationships within maternal families, as well as spatial clustering of genotypes. We find no evidence of frequency-dependent pollen export, suggesting similarity to floral neighbours is more important for female than male fitness. We also find increased mating among plants of the same pigmentation genotypes relative to random mating. Our results indicate that pollinator behaviour causes assortative mating in this population, which would act as a reproductive isolating barrier between *A. m. pseudomajus* and *A. m. striatum*.

4.2 Introduction

Reproductive success is a key determinant of an individual's fitness, and variation in fitness is the driving force behind evolutionary change (Darwin, 1859). The specific patterns of mating among individuals of different genotypes can have important implications for evolutionary outcomes. Assortative mating, where individuals preferentially mate with individuals similar to themselves, may increase the risk of inbreeding depression, but is an important element in the evolution of reproductive isolation (Darwin, 1876; Grant, 1949; Johnson, 1982). Moreover, the distances over which individuals disperse to find mates determines the rate of spread of alleles, with important consequences for the spread of alleles through a population (Wright, 1943). Reconstructing mating events can therefore be a powerful tool in the investigation of evolutionary processes in natural populations.

Many studies have used correlative or indirect methods to investigate reproductive success and mating patterns in plants. Direct observations of pollinator behaviour can identify particular traits or demographic conditions which covary with pollinator visitation (e.g. Waser and Price, 1981; Stanton, 1987a; Harder and Johnson, 2009) or fruit/seed set (Meléndez-Ackerman and Campbell, 1998; Tastard et al., 2012). Observations can also reveal floral constancy to particular floral forms by pollinators (Bateman, 1951; Waser, 1986), whilst dispersal can be inferred from tracking the movement of marked pollen or seeds (Levin and Kerster, 1974). However, these indirect approaches make assumptions which may be violated in reality. Firstly, a pollinator may fail to deliver pollen from or export pollen to a conspecific, so that visitation does not lead to mating (Kobayashi et al., 1999). Secondly, hermaphrodites receive on average half their fitness contribution through female function and half through male function, yet seed set only considers the former. Studies of selection on male fitness are much harder to carry out, and there have been far fewer of them (reviewed in Stinchcombe, 2014). Direct reconstructions of mating events are needed to clarify components of mating success and patterns of mating among individuals in natural populations.

Variable molecular markers can be used to reconstruct genealogical relationships, and hence directly infer mating events and male fitness (Pemberton, 2008; Stinchcombe, 2014). Following the quantitative genetics framework described by Lande and Arnold (1983), early studies assigned paternity to seedlings from open pollinated arrays, and attempted to directly correlate mating events with phenotypes (e.g. Schoen and Stewart, 1986; Devlin

and Ellstrand, 1990). However, because the paternal phenotypes are inferred rather than directly observed, this violates the underlying assumptions of linear regression (Smouse et al., 1999). Furthermore, open-pollinated families consist of one or more groups of individuals sharing a single father. For an animal-pollinated plant with many possible pollen donors, these full-sibling groups probably represent a single mating event, and it is therefore necessary to identify full siblings and assign paternity to the whole family. Methods exist for assigning individuals to full sibships (Wang and Santure, 2009; this volume, chapter 3), but we are not aware of any studies which have applied them to seedling arrays. Although genealogical inference has enormous potential for investigating mating patterns, it is essential to account for uncertainty in the sibship structure in a family, and in the paternity of each sibship.

Distances between pollen donor and maternal plants reflect an underlying distribution of distances, referred to as the pollen dispersal kernel. Some studies have assumed that medium- and long-range dispersal is negligible (Levin and Kerster, 1974), or else that distances decline exponentially (Adams and Birkes, 1991; Burczyk et al., 2002). More recently studies have found that 'fat-tailed' (right-skewed or leptokurtic) dispersal kernels, characterised by rare long-distance events, are common in plants (e.g. Austerlitz et al., 2004; Oddou-Muratorio et al., 2005). Such rare events can be crucial for the biology of wild populations, such as genetic rescue of inbred populations (Dick, 2001), or rapid reforestation after the last ice age (Clark, 1998). Failing to allow for the shape of the dispersal kernel can seriously bias biological inference.

Antirrhinum majus is a short-lived perennial native to the Pyrenees, which shows a striking polymorphism for flower colour in the wild. Where the ranges of the magenta-flowered *A. m. pseudomajus* and yellow-flowered *A. m. striatum* meet they form a narrow hybrid zone, where alleles at two pigmentation loci recombine to give six pigmentation phenotypes (Whibley et al., 2006). The sharp transition in flower colour indicates strong selection on pigmentation loci. Previous work found no evidence for pollinator preference for particular phenotypes, but demonstrated increased pollinator visitation to locally common phenotypes (this volume, chapter 2). This is reflected in increased female fitness for common phenotypes (Tastard et al., 2012), but nothing is known about the male component of fitness. Näive bumblebees in experimental arrays also showed constancy for flower colour, which was sufficient to explain a deficit of hybrids in the resulting progeny

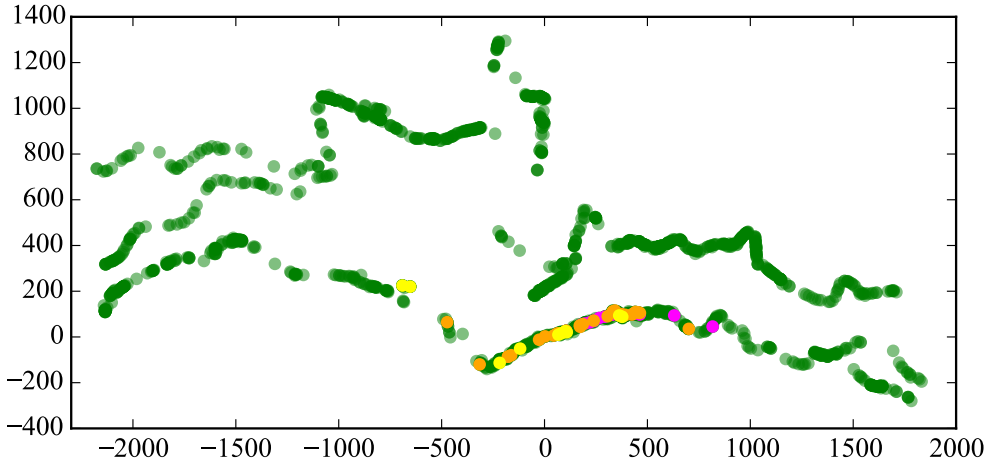


Figure 4.1: Map of sampled plants in 2012. Green points are individuals sampled as part of the population survey. Remaining points are maternal plants used in the paternity study. Magenta points are *A. m. pseudomajus*-like, yellow points are *A. m. striatum*-like and orange points are hybrids. Scale bars show distances in metres.

(this volume, chapter 2). In this study we use paternity analysis of wild-pollinated seeds to directly estimate the effect of pollinator behaviour on the hybrid zone. Specifically we ask: firstly, how important is the local floral neighbourhood for male fitness? Second, do we observe assortative mating for flower colour genotypes? Finally, what is the shape of the dispersal kernel? Our analyses are carried out in a framework which allows us to account for uncertainty in about paternity relationships among offspring, and we fit a model of pollen dispersal which allows for long range migration.

4.3 Materials and Methods

4.3.1 Study population

The study site and population survey is at the boundary between the yellow-flowered *Antirrhinum majus striatum* and the magenta-flowered *A. m. pseudomajus* and is described in detail in chapter 2. These populations form a narrow hybrid zone over ∼1km, where alleles at two floral pigmentation

loci controlling anthocyanin (*Rosea)* and yellow aurone *(Sulfurea)* distribution recombine to give striking transgressive phenotypes (Wheldale, 1907; Baur, 1924; Whibley et al., 2006; figure 1.2). *ROS/ROS* plants have intense magenta pigmentation, whilst *ros/ros* plants have none, and heterozygotes are intermediate. The *SULF* allele is fully dominant over *sulf*, and *SULF/SULF* and *SULF/sulf* plants cannot be distinguished by human eyes, and will henceforth be referred to collectively as *SULF/+*. *sulf/sulf* plant are yellow throughout the face of the flower, whilst *SULF/+* plants express aurones only in a small patch at the centre of the flower.

We surveyed as many flowering plants as we could find in June and July of 2012 (n=2128, figure 4.1), and collected information on flower number and location using a Trimble GeoXT datalogger. *Antirrhinum* grows in disturbed habitat such as roadsides and railways, so it is likely that we sampled almost all of the plants which could have contributed pollen to the sample of mothers. We also collected two to three leaves for DNA extraction and dried these in silica gel (Fischer Scientific). Following the scoring system developed by Whibley et al. (2006) we assigned plants to one of six flower phenotypes based on their inferred genotypes at *Rosea* and *Sulfurea*. As described in chapter 2, we also calculated the density of neighbouring *Antirrhinum* plants around each mother by applying a gaussian function to the distances from a mother to every other plant in the population, and summing over the resulting vector. We used a standard deviation for the gaussian function of 56m, because previous analysis showed that this width best explained variance in pollination rate (chapter 2, figure 2.1). We then used these density estimates to calculate an index of local floral frequency as the density of neighbouring plants of the same *Rosea/Sulfurea* phenotype as the focal plant divided by the density of all neighbours. A plant's floral frequency therefore ranges between one if all neighbouring plants are the same flower colour phenotype as itself, or zero if none of the neighbouring plants are of the same flower colour phenotype as itself.

4.3.2 Progeny array genotyping

In August 2012 we collected a single, mature, wild-pollinated fruit from each of 96 mothers (figure 4.1). In order to minimise disturbance to the population we only sampled from plants which had set a minimum of five mature fruits. We germinated 50-70 seeds from each of 59 of the 96 maternal as described in section 2.3.3, except that we transferred seedlings to 5cm plug

	SULF/+	*sulf/sulf*
ROS/ROS	17	9
ROS/ros	8	9
ros/ros	6	10

Table 4.1: Numbers of maternal plants of each two-locus floral pigmentation genotype.

trays rather than pots. These mothers were chosen to represent an even sample of pigmentation genotypes, spread as evenly as possible across the core of the hybrid zone where hybrids are most dense (table 4.1, figure 4.1). We then grew seedlings in a greenhouse under Sylvania GroLux lights on a 16-hour cycle until at least two pairs of true leaves had appeared. We were able to collect tissue for a total of 1127 seedlings, with between three and 35 individuals per family (mean=19.1). For seedling tissue we transferred approximately $1cm^2$ of fresh tissue to 96-well DNA-extraction plates (LGC Genomics, Berlin) and allowed tissue to dry using the sample bag and silica gel provided. For parental tissue from the hybrid zone we transferred approximately $1cm^2$ silica-dried tissue to the same plates. DNA extractions of the plated tissue samples were carried out by LGC Genomics.

We genotyped tissue samples at 70 SNPs by KASPR sequencing (LGC Genomics). These SNPs are a subsample of a panel used for a wider survey of the hybrid zone (David Field, unpublished data). We identified candidate SNPs from whole genome sequence data (Illumina HiSeq) of eight pools of 50 individuals across a transect through the hybrid zone. The total SNP panel is a mixture of diagnostic (showing a gradient in allele frequency) and parentage (with as even a gradient in allele frequency as possible) SNPs. For parentage loci we chose only biallelic loci with a minor allele frequency $\geq$0.3 in each of inner four pools closest to the centre of the cline, selected to maximise mapping distance between loci. Diagnostic SNPs were either linked to pigmentation loci, or else showed sharp clines across the hybrid zone. To select a subsample of SNPs for this study, we selected markers which were at least 2cM apart to maximise mapping distance between individuals.

Unfortunately, DNA quality for the seedlings was highly variable and 29 SNPs did not amplify reliably. For the remaining 41 SNPs, overall dropout rates per diploid SNP were 2.5% in the offspring and 1.5% in the parents. Minor allele frequencies in the 2128 parents ranged from 0.22 to 0.50.

4.3.3 Sibship and paternity assignment

Maternal families consist of a set of individuals sharing a known mother, divided into subgroups of full siblings who share a father. Failing to account for this family structure can seriously bias inferences about mating patterns. We applied the method described in chapter 3 to sample possible structures for each maternal family using scripts written in Python. This method builds a matrix of probabilities that each candidate father is the true father of each offspring individual (Meagher, 1986; Devlin et al., 1988), and then uses this to construct a bifurcating dendrogram of relatedness between individuals. Bisecting this dendrogram at different points partitions individuals into N different sets of full sibships, where N is the total number of individuals in the half sibling array (figure 3.3), and the relative support for each partition structure can be calculated.

Population sampling at the time of peak of flowering is intense, especially in the area where we collected seeds, so it is likely that most pollen donors were sampled. We allowed for the possibility of a father having been missed by including the probability that an offspring's genotype is drawn from population allele frequencies, multiplied by the prior probability of that father being missing (Nielsen et al., 2001; Appendix to chapter 3). We used a fixed prior on the probability that the true father had been missed of 5%.

We used empirical estimates of allele frequencies from the parental population and a genotype mis-score rate of 1.5% per haploid SNP based on duplicate genotyping of the parents (David Field, unpublished data). We considered each of the 2128 plants sampled in 2012 as a potential father. However, since *A. majus* is self-incompatible in the wild (Xue et al., 1996) we set the probabilities that the mother is also the father of an offspring to zero. Since sibship assignment is prone to 'splinter' individuals into singleton families (chapter 3) we applied a geometric prior on sibship size with p=0.2 to reflect a prior expectation of fewer, larger sibships. We stored all possible structures for each family from their dendrogram of relatedness, along with the posterior probability for each.

4.3.4 Sampling paternal distributions

A key aim of this study is to assess which phenotypes or genotypes explain mating for our sample of maternal plants. Since the pool of candidate fathers is much larger than our sample of mothers, we inevitably miss most of the

pollen transfer in the population. As such, this dataset does not allow us to get reliable estimates of individual fertilities (c.f. Smouse et al., 1999). Instead we examine the probability that mothers receive pollen from a donor of a given phenotype or genotype, relative to the expectation under random mating proportional to spatial distance, summed over all donors of that type.

We used a sampling approach to draw two distributions of possible fathers to compare "credible" and "random" (analogous to observed versus expected) mating events. Credible fathers represent samples of mating events consistent with genotype data. Random fathers represent the expected null distribution if mating is explained solely by geography. We first drew a family structure for each maternal family proportionally to its posterior probability, and multiplied individual paternity likelihoods for each father over every individual in the full sibship (Wang, 2004; this volume, chapter 3). For each full sibship we then sampled a credible father in proportion to the vector of probabilities of paternity for that sibship based on genetic data. We then sampled a random unrelated candidate father proportionally to his distance from the mother, based on the dispersal kernel (see below for details). We repeated this procedure 1000 times for each maternal family, gaining a sample of 259,243 credible or random mating events. This approach allowed us to approximate the posterior distribution of individuals contributing to the pollen pool, whilst simultaneously accounting for uncertainty in sibship structure, paternity, different contributions among maternal plants, and spatial structure among candidate fathers.

The matrix of probabilities of paternity included the probability that the true father of a sibship had not been sampled. The sample of credible fathers therefore accounts for the possibility of missing fathers. Since no information on paternal phenotype or position can be gained from missing fathers, these were excluded from subsequent inference about paternal phenotypes and pollen dispersal. This lends greater weight to full sibships for whom fathers can be identified more clearly.

To assess the effect of floral frequency on pollen export we compared the mean frequency of credible and random fathers. We examined assortative mating among pigmentation genotypes by comparing the ratio of probabilities that a mother of genotype x receives pollen from a father of genotype y for credible and random fathers. When this ratio is one, observed pollen receipt no greater than would be expected by chance. Ratios greater than one indicate increased pollen receipt relative to random mating, and vice versa. To allow for comparison with other studies we also calculated the index of

reproductive isolation as $RI = 1 - 2\,(1 - \Pr(C))$, where $\Pr(C)$ is the probability of receiving pollen from a father of the same genotype (Bateman, 1951; Sobel and Chen, 2014). This ranges from 1 for complete assortative mating to -1 for complete disassortative mating.

4.3.5 Dispersal kernel

We applied a generalised gaussian curve to inter-mate distances to characterise the distribution of pollen dispersal. The generalised gaussian distribution is defined by a scale parameter a and kurtosis parameter b, and has the form:

$$\Pr(d_{ij}|a,b) = \frac{b}{2\pi a \Gamma(1/b)} \exp\left[-\left(\frac{d_{ij}}{a}\right)^b\right]$$

where d_{ij} is the absolute distance between individuals i and j, and $\Gamma()$ is the gamma function. Mean dispersal is given by the second moment of this distribution:

$$\frac{\alpha^2 \Gamma(3/b)}{\Gamma(1/b)}.$$

This distribution has two advantages for modelling plant dispersal (Clark, 1998). Firstly, it generalises standard exponential family distributions; when $b = 2$ it corresponds to a gaussian function with $a = \sigma\sqrt{2}$, and when $b = 1$ it is equivalent to an exponential distribution with rate a (Austerlitz et al., 2004). Moreover, it can explicitly account for rare long-distance migration events via the kurtosis parameter. Values of $b < 1$ indicate that most dispersal is over short distances, but with an excess of long-distance events relative to exponential decay.

Since no closed-form maximum-likelihood estimator exists, we fitted the dispersal kernel to the distances from mothers to credible and random fathers by sequentially maximising a and b. Starting with $b = 1$, we found the value of a between 1 and 2000m with the highest log likelihood, then used this value to maximise b between 0.1 and 3. We repeated this procedure until the likelihood could not be improved. To estimate dispersal relative to the direction of the hybrid zone we fitted kernels to the East-West distances between fathers and mothers.

To gain a dispersal function for sampling random fathers we also attempted to fit a curve for euclidean distances. However, because plants are

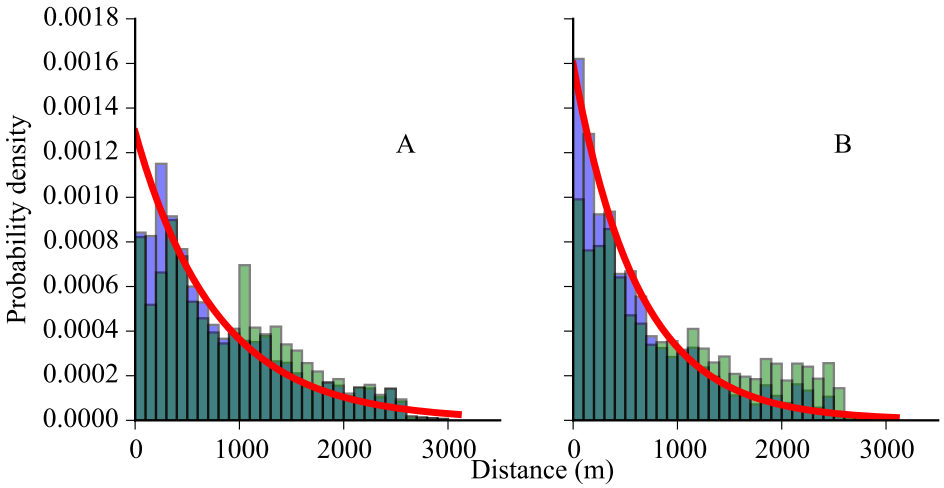


Figure 4.2: Distances between maternal plants and candidate fathers. (A) shows the distribution of distances between mothers and 2127 candidate fathers if distance incorporates East-West, North-South and altitude directions (green), or only East-West and altitude (blue). Note the hump in distances at around 300m. The red curve is the best fitting curve to the latter dataset. (B) shows the distribution of distances to mothers in the East-West direction only. Blue bars show the distribution for the sample of credible fathers, green bars show the distribution for all 2127 candidates, and the red curve is dispersal kernel used to draw random fathers. Where green and blue bars overlap, the colour is darker green.

clustered along two roads, incorporating North-South distance causes this distribution to be bi- and possibly trimodal, which renders a unimodal dispersal kernel meaningless (figure 4.2A). As a compromise, we instead used only the East-West and altitude directions, applying a generalised gaussian distribution to distances with $a = 0.99$ and $b = 786$m (figure 4.2B). Since altitude is strongly correlated with the North-South axis, this allows for this spatial variation whilst still permitting kernel fitting to a unimodal distribution.

4.3.6 Confidence intervals

To estimate confidence intervals (CIs) for the distributions of relative pollen receipt we drew 1000 subsamples without replacement from the sample of mating events and calculated parameter estimates for each subsample (Politis and Romano, 1994). 95% confidence intervals are the 97.5% and 2.5% quantiles of the distributions across subsamples. To calculate p-values for assortative mating pattern we use the proportion of the subsample distribution which overlapped 1. To calculate p-values for frequency we used the proportion of subsamples for which the mean for credible fathers was greater than the mean for random fathers.

4.4 Results

Mean East-West distance between mothers and credible fathers was 645m (CIs=605, 685m; figure 4.2). The best fitting dispersal kernel was approximately exponential with a=626m (CIs=513, 685) and b=0.98 (CIs=0.88, 1.02). We found no difference in the average frequency of credible and random fathers (means=0.483, 0.488; p=0.387), and credible fathers seems to be a random draw from the population distribution (figure 4.3).

The probability of mating increased with the number of *Rosea* alleles shared between maternal plants and pollen donors (figure 4.4). *ROS/ROS* and *ros/ros* maternal plants received significantly reduced pollen from their respective opposing homozygous genotypes than would be expected by chance (p=0.0066, <0.0001). *ros/ros* mothers received increased pollen from *ros/ros* and *ROS/ros* donors (p=0.0061, 0.0017) and had a mean RI index of 0.469. *ROS/ROS* mothers were no more likely to receive pollen from *ROS/ROS* or *ROS/ros* plants than expected by chance (p=0.0931, 0.1608), and had a

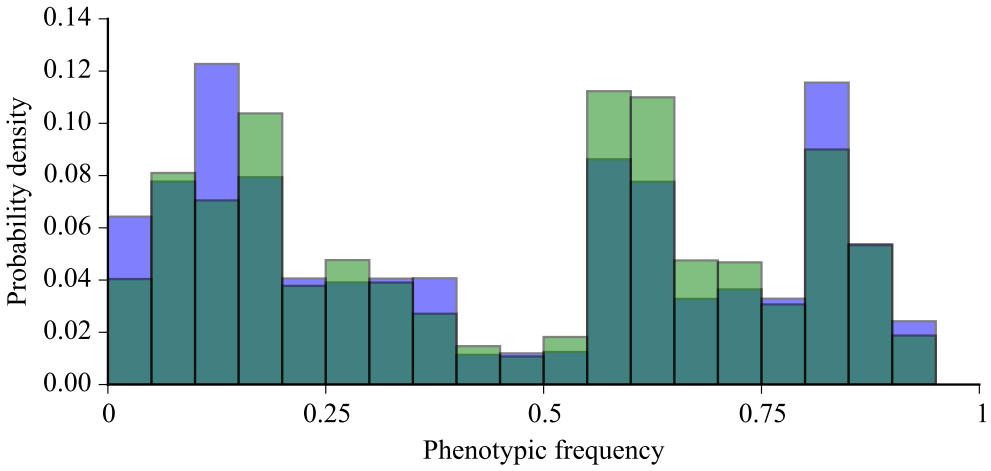


Figure 4.3: Histogram of floral frequencies for credible fathers (blue) and random fathers (green). Where bars overlap the colour is darker green.

mean RI index of 0.633.

We observed non-random mating patterns for *Sulfurea* genotypes only when the maternal plant was *sulf/sulf*. *SULF/+* plants were no more or less likely to receive pollen from *SULF/+* or *sulf/sulf* donors than would be expected under random mating (p= 0.2836, 0.2838). In contrast, *sulf/sulf* mothers received significantly reduced pollen from *SULF/+* donors (p<0.0001) but increased pollen from *sulf/sulf* donors (p<0.0001). In accordance with these results, mean RI indices were -0.025 for *SULF/+* mothers and for 0.311 for *sulf/sulf* mothers.

4.5 Discussion

We used a combination of sibship and paternity assignment to reconstruct mating patterns in a hybrid zone of *Antirrhinum majus*. Contrary to our expectation we found no evidence that locally common flower-colour phenotypes have increased pollen export. Nevertheless we found a clear pattern of assortative mating for floral pigmentation loci, with more mating among similar genotypes than would be expected by chance.

Unfortunately, due to issues with DNA-quality, many SNP markers failed

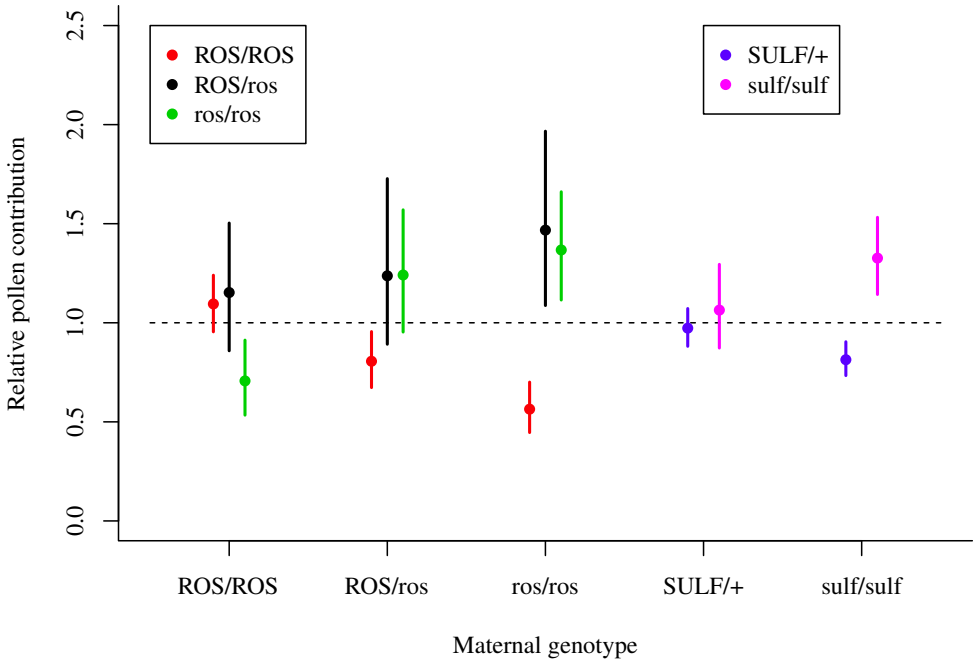


Figure 4.4: Mating patterns for *Rosea* and *Sulfurea* genotypes. Points show the probability that a stigma receives pollen from a credible father of a particular genotype, relative to a random father for different maternal genotypes.

to amplify reliably, and we were not able to resolve relationships as well as we would have liked. The difficulty of identifying a single true parent increases with the number of alternative candidates because there is an increased risk that other candidates are more closely related to the offspring by chance. This is a particular challenge for this dataset because since the number of candidate fathers is around ten-fold greater than in any similar studies. Our sample compared 2127 candidates, but we are not aware of any studies reporting samples of more than 286 candidates (Hodgins and Barrett, 2008). Reanalysis of this dataset with extra markers would improve resolution of genealogical relationships.

The apparent lack of a relationship between paternal floral frequency and pollen export contrasts with previous work showing that frequency predicts both pollinator visitation and fruit set (this volume, chapter 2; Tastard et al., 2012). Inspired by work on animals (Bateman, 1948), it has been hypothesised that selection on floral traits acts primarily on male function (Bell, 1985; Burd, 1994), which would directly contradict our results. This may be due to a lack of SNPs, and hence statistical power. However, meta-analysis of empirical evidence does not support the male-function hypothesis (Ashman and Morgan, 2004). Paternity studies have often found either no selection on male function, or variable environment-dependent selection (e.g. Kobayashi et al., 1999; Smouse et al., 1999; Wright and Meagher, 2004; Hodgins and Barrett, 2008). It may well be that floral frequency is important primarily for pollen receipt in *A. majus,* while pollen export is so diffuse that selection is weak. This is consistent with the observation of widespread pollen limitation in this population (Tastard et al., 2012). Furthermore, we only examined paternity in a single year, but both pollinator visitation and seed set varied among years (this volume, chapter 2; Tastard et al., 2012), so there may be temporal variation in the strength of selection on floral frequency. Thus, while we have no evidence in favour of frequency-dependent selection on male function, it is premature to rule it out entirely.

Our dataset reveals a distinct pattern of assortative mating for flower colour. The probability of pollen export to *Rosea* and *sulf/sulf* homozygotes increased with the number of shared pigmentation alleles between mother and father (figure 4.4). The lack of assortment for *SULF/+* mothers is probably influenced by interactions with *Rosea* genotype. We cannot rule out that there are cryptic differences between *SULF/SULF* and *SULF/sulf* plants, but it is not obvious why that should be. Because we compared paternal genotype frequencies to expectations under random mating and a dispersal

kernel, these findings cannot be explained as the spurious results of spatial clustering. Andalo et al. (2010) found no evidence for postzygotic barriers between *A. m. striatum* and *A. m. pseudomajus*, so genetic incompatibilities are unlikely to explain these patterns, although it is possible that late-generation hybrid breakdown occurs. The observed indices of reproductive isolation for *Rosea* (0.63) and especially *Sulfurea* (0.47) are similar to that found for hybrid deficit in experimental arrays with naïve pollinators (0.42; this volume, chapter 2). These lines of evidence suggest that floral constancy by pollinators drive the patterns of assortative mating observed in the hybrid zone.

The estimate of mean pollen dispersal of 645m is surprisingly far, and two to three times greater than estimates for four tree species given by Austerlitz et al. (2004). Bumblebee workers and queens do forage over such distances(Osborne et al., 2008; Lepais et al., 2010; Hagen et al., 2011), but this is likely to be an overestimate for two reasons. Where the true father has not been genotyped, the probability distribution of the remaining candidates will be close uniform, and any one could be drawn as a credible father (in proportion to the prior probability that the father has been sampled). Moreover, the modest power of the SNP dataset, coupled with the large number of candidate fathers means that the posterior probability of drawing the correct father will be reduced. Both of these factors will cause unrelated candidates to be drawn as credible fathers, representing essentially a random draw from the sample of candidate males. This will in turn cause the distribution of inferred to dispersal events to resemble the distribution of distances between mothers and all possible candidate males (i.e. the blue and green bars respectively in figure 4.2B). It is encouraging that the distances to credible fathers is shifted to the left-hand side of figure 4.2B, indicating that the sample of credible fathers are not simply a random draw from the population, but actually reflect some of the real mating patterns in the hybrid zone, albeit with considerable noise. Nevertheless, more genetic markers are needed before robust conclusions can be drawn about pollen dispersal in this system.

Our results have shown that floral constancy by pollinators does indeed lead to assortative mating in the *A. majus* hybrid zone, which will act as a reproductive isolating barrier between *A. m. pseudomajus* and *A. m. striatum.* Given this non-independence of parental genotypes it should be noted that the mating patterns we have observed are conditional on the set of maternal plants sampled. Nevertheless we endeavoured to sample as even

a sample of maternal plants as possible with regards to pigmentation and spatial arrangement (table 4.1, figure 4.1), so our results probably do reflect the range of mating patterns experienced by plants across the hybrid zone.

4.6 Collaborator contributions

Contributions to the population survey are described in chapter 2. Seeds were collected by David Field and myself in August 2012. David Field developed the SNPs and arranged genotyping with LGC Genomics. Plating parental tissue for DNA extraction was performed by David Field and Sylvia Rebel. David Field also estimated the SNP genotype error rates based on repeat genotype data.

Chapter 5

Repeated gains in yellow and anthocyanin pigmentation in flower colour transitions in the Antirrhineae

5.1 Abstract

Angiosperms display remarkable diversity in flower colour, implying that transitions between pigmentation phenotypes must have been common. Despite progress in understanding transitions between anthocyanin (blue, purple, pink or red) and unpigmented (white) flowers, little is known about the evolutionary processes driving flower-colour transitions in lineages with both yellow and anthocyanin pigmented flowers. We use phylogenetic comparative analyses of 178 species from the tribe Antirrhineae to investigate how anthocyanin and yellow pigments interact to shape phenotypic diversity. In contrast to previous studies we found a bias towards transitions involving a gain in pigmentation, but taxa with both yellow and anthocyanin pigmentation are extremely rare. This indicates that in the long-term selection tends to promote the spread of single-pigmented morphs We also find an overall greater flux between anthocyanin and white phenotypes throughout the history of this group, which correlates with high levels of extant polymorphism. Our findings show that interactions between anthocyanin and yellow pigments constrain the breadth of potential floral diversity observed in nature.

In particular, they show that evolutionary forces act to promote the spread of single-pigmented phenotypes across the Antirrhineae phylogeny. Furthermore the correlation between transition rates and polymorphism suggest that the forces causing and maintaining variance in the short term reflect evolutionary processes on longer time scales.

5.2 Introduction

Colours in nature are important cues for organisms to signal warnings or rewards. In angiosperms, flower pigmentation is important for pollinator attraction and is linked to a suite of non-pollinator-related traits (Faegri and Van der Pijl, 1966; Strauss and Whittall, 2006). Transitions in flower colour across a phylogeny are common in many plant lineages, which allows us to draw conclusions about their mechanism and consequences from many replicated evolutionary events (Rausher, 2008; Streisfeld and Rausher, 2011; Wessinger and Rausher, 2012). Moreover, mutations causing flower colour changes are abundant and conspicuous, and the underlying synthetic pathways are well understood (Grotewold, 2006). Flower colour therefore provides an excellent model for the investigation of evolutionary change because we can link molecular genetic changes to the ecology and demography of the organism.

Floral pigments fall into only a handful of molecular families conserved across the angiosperms (Rausher, 2006). To date most of what we know about flower colour transitions has focussed on the gain or loss of one or more anthocyanin pigments, which confer red, blue, pink and purple colours (Grotewold, 2006; Rausher, 2008; Wessinger and Rausher, 2012). In the wild, mutations at a single transcription factor regulating structural anthocyanin synthesis enzymes have been shown to be sufficient to effect a gain or loss of anthocyanin pigmentation in *Petunia* (Quattrocchio et al., 1999), *Antirrhinum* (Schwinn et al., 2006), *Aquilegia* (Whittall et al., 2006) and *Mimulus* (Cooley et al., 2011; Streisfeld et al., 2013), although additional mutations may occur later (Zufall and Rausher, 2004). However, many plant lineages also include species with yellow flowers, with pigments derived from the carotenoid, aurone or betalainin pathways, which are distinct from anthocyanins (Grotewold, 2006). Transitions involving the gain or loss of yellow pigments have been less well studied, but studies on carotenoid loss in Brassica and Chrysanthenum point to loss-of-function mutations at loci in-

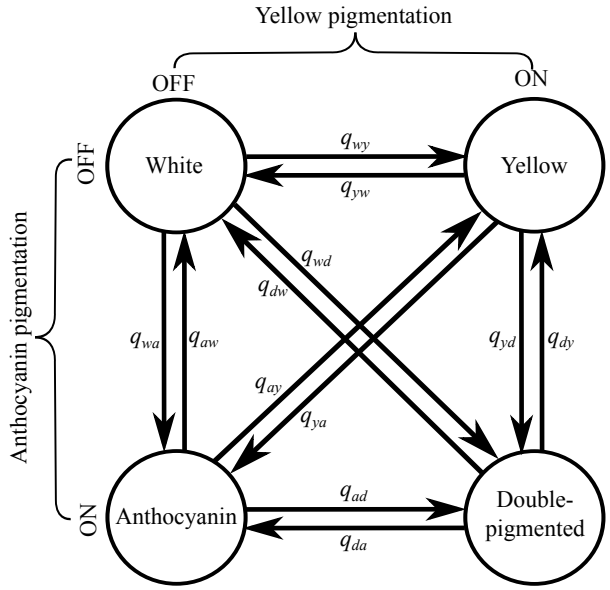


Figure 5.1: Model for the pigmentation state space and possible transitions between them. Interaction between anthocyanin and yellow pigment pathway can give rise to four phenotypes. The transition rate from phenotype i to phenotype j across the phylogeny between two phenotypes is labelled q_{ij}.

volved in the downstream degradation of pigments to colourless compounds (Ohmiya et al., 2006; Zhang et al., 2015). A pattern common to all floral colour transitions examined so far is that a mutation at a single locus is sufficient to cause a shift in flower colour.

We currently know little about the evolution of flower colour when both anthocyanin and yellow pigments are present. Different combinations of yellow and anthocyanin pigments give rise to four colour phenotypes (figure 5.1). A floral tissue may produce neither pigment (white), one pigment class only, or both yellow and anthocyanin pigments together, conferring an orange, or in some cases red phenotype (Hackbarth et al., 1942; Stanton, 1987b; Cooley et al., 2011; Streisfeld et al., 2013). The addition of a second pigment pathway makes transitions between yellow and anthocyanin pigmentation fundamentally different from those previously studied because they imply changes at a minimum of two loci; one pathway must be turned off and the other turned on. Alternatively, apparent "dual" transitions might occur if a single mutation influences both pigment classes at once (q_{wd} and q_{dw}, fig-

ure 5.1) or if one pigment is masked by the other, but becomes visible once that pigment is removed (q_{ay} and q_{ya}; figure 5.1). Groups of plants with both pigments provide considerable scope for interactions (e.g. epistasis and pleiotropy) that may constrain evolutionary transitions.

In this study we examine the history of transitions in flower colour in the tribe Antirrhineae. This group is comprised of approximately 370 mostly short-lived perennial herbs distributed throughout Eurasia and North America (Sutton, 1988). Floral pigmentation is present in 90% of species, and pigment morphs are well distributed among genera (Sutton, 1988; figure 5.2). The genetic basis of pigmentation in the model snapdragon *Antirrhinum majus* has been studied for over a century, and is known to follow the two-dimensional model described above and in figure 5.1 (Wheldale, 1907; Baur, 1924; Whibley et al., 2006). Flower colour in *A. majus* is derived primarily from the magenta anthocyanin cyanidin and yellow aurone pigments (Geissman et al., 1954; Ono et al., 2006; Schwinn et al., 2006), and crosses between species show that homologous loci are present throughout the old-world *Antirrhinum* (Hackbarth et al., 1942). Studies beyond *Antirrhinum* have been limited, but work on *Linaria*, the largest genus in the tribe, have recovered similar pigments and genetic architectures (Tjebbes, 1929; Harborne, 1966; Valdés, 1970). It should be noted that aurones and anthocyanins are both derived from chalcone pigments, whilst carotenoids are not. There is therefore greater potential for pleiotropy and substrate competition between these pathways (Ono et al., 2006), which might not be these case for systems with carotenoid pigmentation. Nevertheless, the abundant variation in pigmentation phenotypes in the Antirrhineae make this tribe a promising testing ground for investigations of transitions between yellow and anthocyanin-pigmented floral phenotypes. Only one study has examined flower colour transitions in a phylogenetic context for the Antirrhineae. This study revealed a bias towards gain of anthocyanins (Smith and Goldberg, 2015). However, they considered only a section of Antirrhinum, and did not consider yellow pigmentation.

Here we present the first attempt to examine patterns of flower colour transitions when two floral pigments underlie colour variation. We use phylogenetic comparative analyses of 169 species from the Antirrhineae to investigate the evolutionary processes acting on anthocyanin and yellow pigmentation. We reconstruct phylogenetic relationships and then estimate transition rates between colour phenotypes. Our results suggest that long-term evolutionary transitions occur at one pathway at a time, especially between

Figure 5.2: Majority-rule consensus tree of Antirrhineae taxa for which molecular data are available. Circles indicate floral colour phenotype. Polymorphic taxa have more than one circle.

anthocyanin-pigmented and unpigmented taxa. These results provide novel insights into the role of interactions between pigments in constraining the possible evolutionary pathways between flower colour phenotypes.

5.3 Materials and Methods

5.3.1 Flower colour

We collected data on flower colour from classical taxonomic literature (Chavannes, 1833; Munz, 1926; Pennell, 1947; Rothmaler, 1956; Speta, 1980; Elisens, 1985; Fernandéz Casos, 1988; Sutton, 1988; Thompson, 1988; Güemes, 1994; Whibley et al., 2006) on all species given by Sutton (1988). The corolla of Antirrhinneae typically have a "major" colour, but often have some secondary pigmentation, such as purple veins or a yellow palate. We recorded the primary colour throughout the face of the flower described by the authors as either white, yellow, anthocyanin pigmented, or double pigmented. Since it was not possible to differentiate red, blue, pink or purple clearly from many descriptions, we grouped these together as anthocyanin pigmented. We classified ambiguous descriptions of pale flowers such as "whitish-pink" or "whitish-yellow" as anthocyanin or yellow respectively, since these cases represent low levels of pigmentation. In some cases, such as species of *Mimulus*, double-pigmented flowers appear red rather than orange (Cooley et al., 2011; Streisfeld et al., 2013). We could not distinguish single- and double-pigmented red flowers from taxonomic descriptions, and scored these as anthocyanin-pigmented phenotypes. However, since only 6.7% of the taxa included in the phylogeny (below) are red, any true double-pigmented red morphs are unlikely to have a large effect on our results.

Colour descriptions in taxonomic literature are often brief and descriptive, but it is not practical to collect spectral reflectance data for such a large, widely dispersed tribe. We took several measures to try and remove ambiguous taxa. Where possible (162 taxa) we cross-referenced flower colour descriptions from multiple authors, and excluded two taxa with contradictory sources. We excluded two taxa with ambiguous descriptions among pigment types (*Linaria pedunculata* and *Linaria albifrons*), as well as three taxa with no information on flower colour (*Linaria paradoxa, Misopates salvagense* and *Sairocarpus elmeri*).

Polymorphic taxa are a persistent conundrum in comparative phylogen-

etic analyses. On one hand, a polymorphism reflects an pigmentation variant which has risen to sufficient frequency to be noted by taxonomists, which may represent an incipient transition. In this sense they provide valuable information. On the other hand, this variation may be transient with little effect on long-term trends. Moreover, polymorphic taxa could introduce a spurious correlation between diversification and transition rates if a particular phenotype appears often in a polymorphism. We found 24 taxa described as polymorphic, and entered a separate entry for each colour. We then ran all subsequent analyses on two datasets: one including all available taxa (polymorphic dataset), and another excluding the 24 polymorphic taxa (polymorphic dataset).

5.3.2 Phylogenetic reconstruction

We reconstructed the phylogenetic relationships for 169 Antirrhineae species using internal transcribed spacer sequences retrieved from GenBank (Benson et al., 2000). Details of sequencing and accession numbers are described in Fernández-Mazuecos et al. (2013). We added duplicate sequences for each colour morph for polymorphic taxa, meaning each morph is represented as a bifurcation at the tip of the tree with branch lengths of zero. For the monomorphic dataset, we pruned polymorphic taxa entirely. We aligned these sequences in MAFFT (Katoh and Standley, 2013) and estimated the phylogeny in MrBayes 3.2.2 (Ronquist and Huelsenbeck, 2003) using the GTR substitution model with gamma-distributed rate variation across sites. After a burn-in of 10,000 generations we ran the Markov chain for 10,000,000 generations. We generated a sample of 1000 trees from the posterior distribution by sampling every 10,000 generations to avoid autocorrelation between trees. Including polymorphic taxa, flower colour data were available for 186 tips, representing approximately half of the tribe.

5.3.3 Joint estimation of transition and diversification rates

We used the R package diversitree to jointly estimate transition rates between states whilst accounting for possible differences in diversification rates among states (R Core Team, 2012; Maddison et al., 2007; FitzJohn et al., 2009; FitzJohn, 2012). We denote the transition rate from phenotype i to phenotype j as q_{ij}, and label white, yellow, anthocyanin-pigmented, and double-

pigmented phenotypes *w, y, a* and *d* respectively (figure 5.1). We compared two models of trait evolution. In the "constrained" model we constrained dual transitions between white and double pigmentation (q_{wd} , q_{dw}) and between anthocyanin and yellow pigmentation (q_{ay}, q_{ya}) to zero. To account for the possibility of dual transitions being very rapid, or for one pigment masking another with low expression levels, we also ran a "full" model with all twelve possible transitions between phenotype allowed. Each model was run for datasets with and without polymorphic taxa.

Transition rate estimates can be biased if differences in speciation (λ) and extinction (μ) rate among states are not accounted for. However, it has recently been shown that these estimates are frequently subject to alarming biases due to correlations with unmeasured characters (Rabosky and Goldberg, 2015). We therefore fitted models which allow for different λ and μ among states so that transition rate estimates are not constrained by these, but confine our interpretation of results to transition rates.

We found this dataset to be sensitive to the hill-climbing algorithm implemented in diversitree, which gave radically different results depending on the starting point. We therefore estimated transition rates among floral phenotypes using the MultiMusse MCMC command in diversitree using an exponential prior with a mean of 0.1, and adjusting for taxon sampling. To account for phylogenetic uncertainty we ran the chain for 100 generations on a single tree, and then swapped to a new tree drawn at random from the sample of trees from MrBayes. We first allowed the chain to explore 100 trees as burn-in, and then stored the final generation of chains for each of 1000 subsequent trees. We estimated the marginal likelihood of the full and constrained models as the harmonic mean of the likelihood of each generation in the Markov chain (Newton and Raftery, 1994). This approach has been criticised because it is insensitive to changes in the prior (Friel and Wyse, 2012), but since we used a non-hierarchical model with a fixed prior, the estimator is appropriate. To test the significance of differences in support we applied a χ^2-test with four degrees of freedom to the ratio of harmonic mean likelihoods for both models, corresponding to the difference in the number of parameters to be estimated under each model.

5.3.4 Asymmetry in rate estimates

Hypothesis testing for differences in rate parameters, such as asymmetry in the direction of transitions, is often approached by comparing the degree of

support for two models using likelihood ratio tests (Pagel, 1994). However, an asymmetry can itself bias the test (Goldberg and Igić, 2008) and requires extensive model exploration, which is impractical for this dataset because of the need to use MCMC. Instead we assess asymmetry in transition rates by comparing the vectors of parameter estimates from the posterior distributions of Markov chain outputs. The posterior probability (pp) that parameter x is greater than parameter y is simply the frequency of generations in which $x_i > y_i$ (Kruschke, 2010).

5.4 Results

We collected floral trait data on 346 monomorphic taxa, of which 27 were white, 144 were anthocyanin-pigmented, 170 were yellow, and two were double pigmented. Of these 19, 70, 63 and one taxa respectively could be included in phylogenetic analyses. We found descriptions of polymorphism in a further 23 taxa, including all four combinations of white, anthocyanin-pigmented and yellow morphs, as well as one yellow/double-pigmented polymorphism (table 5.1). Of these, white/anthocyanin polymorphisms account for more than all other combinations combined. A further three taxa were described by Sutton (1988), but no colour information was given by any author. Yellow and anthocyanin-pigmented taxa are well distributed across the phylogeny (figure 5.2) and frequently occur as sister species. In this sample of taxa, white species are distributed sparsely across the tree and occur very rarely as closely-related species. When only monomorphic taxa were included we found better point estimates for parameters, with narrower, more peaked posterior distributions (figure 5.3, supplementary figure 1). This suggests that phylogenetic signal was stronger when polymorphic taxa were excluded. Bayes factor comparisons for models of trait evolution revealed strong support for the constrained model which excludes anthocyanin-yellow and white-double pigmentation transitions (figure 5.1), when only monomorphic taxa are considered (df=4, $p<0.0001$). When polymorphic taxa are included in the dataset there was no difference in support for either the full or constrained model (df=4, p=0.315).

We found a much clearer phylogenetic signal when only monomorphic taxa were considered, reflected in the narrower and more peaked posterior distributions of parameter estimates for this dataset (figure 5.3, supplementary figure 1). Bayes factor comparisons for models of trait evolution revealed

	total	phylogeny
white/anthocyanin	12	8
white/yellow	3	1
anthocyanin/yellow	5	3
yellow/double pigmented	1	1
white/anthocyanin/yellow	2	1

Table 5.1: Frequencies of colour phenotype combinations in polymorphic taxa in the whole tribe (total) and the taxa for which molecular data are available (phylogeny).

hypothesis	monomorphic		polymorphic	
	full	cons	full	cons
$q_{wa} > q_{aw}$	0.705	0.745	0.871	0.910
$q_{wy} > q_{yw}$	0.685	0.756	0.628	0.689
$q_{ya} > q_{ay}$	0.893	-	0.759	-
$q_{dw} > q_{wd}$	0.854	-	0.877	-
$q_{da} > q_{ad}$	0.972	0.999	0.973	0.993
$q_{dy} > q_{yd}$	0.936	0.974	0.962	0.986
$q_{wa+aw} > q_{wy+yw}$	0.905	0.845	0.919	0.932

Table 5.2: Posterior probabilities of rate asymmetries for monomorphic and polymorphic datasets under the full and constrained models.

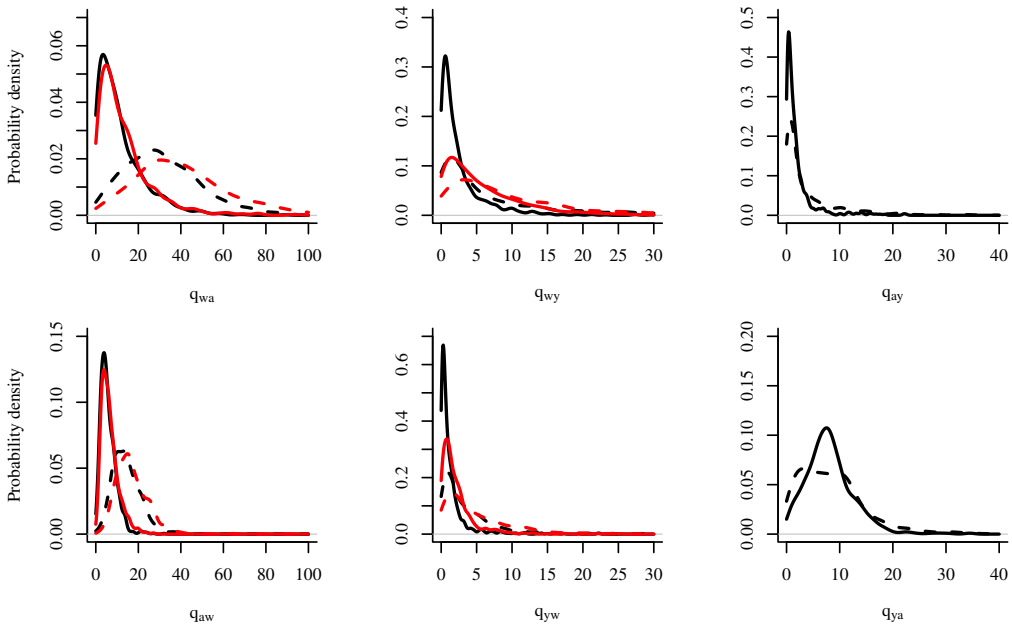


Figure 5.3: Posterior probability distribution of transition rate estimates between white, anthocyanin and yellow taxa. Black lines show the distribution under the full model, and red lines show distributions under the constrained model. Solid lines show distributions for monomorphic taxa only, whilst dashed lines are for datasets inluding polymorphic taxa.

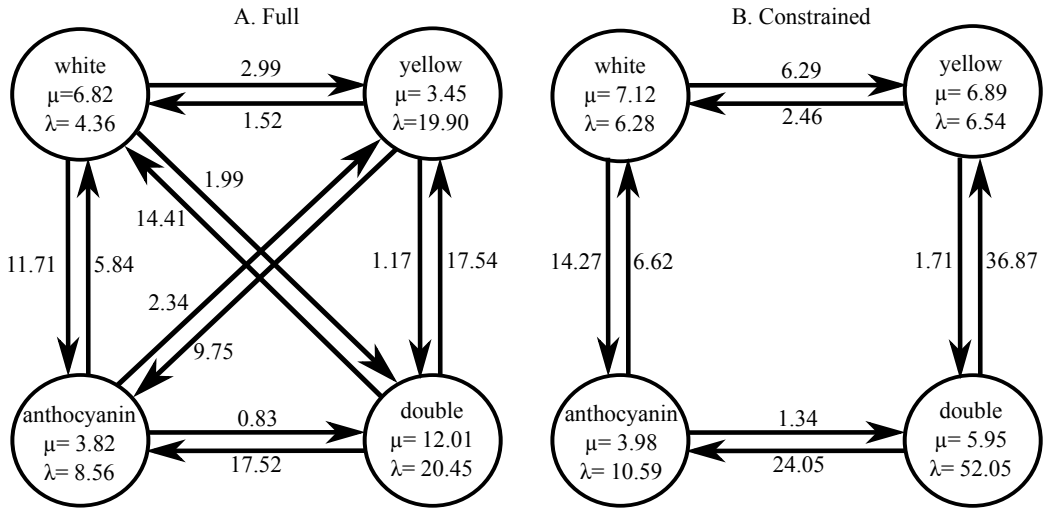


Figure 5.4: Parameter rate point estimates for monomorphic taxa under (A) the full and (B) constrained models. States and transitions correspond to those given in figure 5.1, and point estimates are the means of the posterior distributions.

strong support for the constrained model which excludes anthocyanin-yellow and white-double pigmented transitions, when only monomorphic taxa are considered (df=4, $p<0.0001$). When polymorphic taxa are included in the dataset there was no difference in support for either the full or constrained model (df=4, p=0.315).

We found evidence for asymmetries in transition rates between taxa (figure 5.4, table 5.2). The overall flux between white and anthocyanin pigmentation in both directions was two- to four-fold greater than those between white and yellow pigmentation across models and datasets. Transitions rates away from white to anthocyanin pigmentation were twice those towards white, and this pattern had particularly strong support when polymorphic taxa were included. Similarly we observed a two-fold higher transition rate away from white to yellow than towards white (figures 5.3-5.4), but statistical support for this observation was equivocal (table 5.2). Transitions from yellow to anthocyanin-pigmented phenotypes were around two- to four-fold higher than those towards yellow, but this asymmetry is much weaker when polymorphic taxa are included. Under all models and datasets, transitions away from double pigmentation to all other phenotypes were much greater than

those towards double pigmentation (pp≥0.936). Estimates of speciation and extinction varied, and showed no overall pattern of association with particular states (supplementary figures 2 and 3).

5.5 Discussion

Transitions between flower colour phenotypes have been frequent in the Antirrhineae, and yet the distribution of possible phenotypes is strongly biased towards yellow and anthocyanin-pigmented species. In this study we have used phylogenetic comparative methods and accounted for phylogenetic uncertainty to investigate the patterns in floral colour transitions which have contributed to this heterogeneity. Interestingly, when only monomorphic taxa were considered we found greater support for a model which only allowed substitutions at one pigment pathway at a time. However, when polymorphic taxa were included there was no difference in support for this model or one allowing apparent dual transitions causing apparent substitutions at two pathways simultaneously. This suggests that many of the inferred dual transitions are due to the effect of polymorphism, representing potentially transient events at the tips of the tree. Transitions among monomorphic taxa, in contrast, are more likely to represent completed transitions deeper in the tree, and better reflect long-term evolutionary patterns. This is confirmed by the markedly greater phylogenetic signal for the monomorphic dataset. Greater support for the constrained model implies that transitions in flower colour phenotypes occur primarily via stepwise substitutions at each pigment pathway, with a detectable waiting time in each state. This is consistent with the observation that whilst mutations arise just as frequently in structural and regulatory loci, transitions in flower colour on evolutionary timescales are typically due to mutations in genes regulating specific pigment pathways (Streisfeld and Rausher, 2011; Wessinger and Rausher, 2012). By examining the finer patterns of transitions we can use these data to elucidate possible mechanisms which lead to differences in transition rates.

Low overall rates of transition will hinder the distribution of phenotypes from reaching equilibrium. Our results suggest that the frequency of transitions is not a limiting factor in shifts between unpigmented and single pigmented flowers, consistent with previous observations in *Ipomoea* (Smith et al., 2010). Yellow and anthocyanin-pigmented taxa frequently occur as sister taxa across the phylogeny, and transition rate estimates indicate fre-

quent transitions between these colours via white. In contrast, the paucity of double-pigmented taxa and the very low estimates of q_{wd}, q_{ad} and q_{yd} suggest that transitions to double pigmentation are very rare. It is nevertheless possible that double-pigmented taxa do indeed arise but rapidly go extinct, or that double pigmentation appears only fleetingly before one of the two pigments is lost. We found five species with yellow/anthocyanin polymorphisms where recombination between pigmentation genes could lead to double-pigmented morphs, but this appears not to be common. Double-pigmented flowers do arise in the wild via hybridisation between yellow and magenta populations of *Antirrhinum* majus, but are confined to narrow hybrid zones, suggesting that some selective mechanism prevents their spread (Whibley et al., 2006). It may be that there is a higher cost associated with producing multiple pigments, or that pollinators discriminate against double-pigmented taxa, as has been shown in *Raphanus sativus* (Stanton, 1987b). Thus, although direct evidence is lacking, it is likely that some form of natural selection prevents the spread of double-pigmented colour morphs in the wild, and prevents transitions to double pigmentation across the phylogeny.

We found evidence for asymmetries in the direction of flower colour transitions leading to the gain of pigmentation. Transitions away from white to yellow and especially anthocyanin pigmentation were greater than transitions to unpigmented flowers. Although statistical support was weak, this observation contrasts with other systems which have found a bias in transitions towards white from anthocyanin pigmentation, which are often irreversible (Rausher, 2006; Whittall et al., 2006; Smith et al., 2010). Nevertheless, Cooley et al. (2011) demonstrated parallel gains of anthocyanin pigmentation in two species of the *luteus* group of *Mimulus*. This result also mirrors the findings of Smith and Goldberg (2015), who also found an asymmetry in gains of anthocyanins in a subset of the Antirrhineae, and showed that this asymmetry was stronger than in three other tribes examined. Our data show that this pattern is general across the whole tribe, and holds for yellow pigmentation as well. This asymmetry, coupled with the overall paucity of white taxa suggests that some kind of selective mechanism acts to promote the spread of pigmented morphs in white species. For example, white phenotypes have been shown to be undervisited by pollinators (e.g. Waser and Price, 1981), to be more sensitive to drought stress (Warren and Mackenzie, 2001), and to be associated with greater inbreeding depression (Burdon et al., 1983). However, in other conditions white morphs are preferred by pollinators (Stanton, 1987b), and show higher growth rates than anthocyanin-pigmented morphs

under well-watered conditions (Schemske and Bierzychudek, 2001; Warren and Mackenzie, 2001). The phenotypes observed on the tips of a phylogeny today are the cumulative result of processes acting over long-term evolutionary time scales, and there is ample scope for fluctuating selection to favour different morphs at different times. If selection acts against white morphs for some, but not all of the time, this will tend to increase extinction rates among white species, as well as increase the rate of transitions to pigmented flowers, and could cause the persistence of white taxa at low frequency.

Flower colour polymorphisms often have a simple genetic basis, and molecular studies have shown a strong tendency for transitions to be driven by changes at a single regulatory gene (reviewed in Wessinger and Rausher, 2012). Indeed, anthocyanin pigmentation in *Antirrhinum* majus is controlled by a single *MYB*-transcription factor (Schwinn et al., 2006), and alleles at flower colour loci throughout the genus *Antirrhinum* segregate as Mendelian traits (Wheldale, 1907; Baur, 1924; Hackbarth et al., 1942), so this is at least partially true in the Antirrhineae. The greater support for the constrained model is consistent with a model of repeated substitutions at single loci influencing the anthocyanin and aurone pathways independently. The stepwise model of pigment evolution also suggests that although anthocyanins and aurones are both derived from chalcones (Ono et al., 2006), mutations tend to be specific to a single pathway, and there is little competition for substrates. For Mendelian traits in a monomorphic population (i.e. with no standing variation), the major constraint on adaptation is the waiting time for a new mutation to arise. Since there are many more ways for mutation to cause a loss-of-function than gain-of-function mutation, one would expect a bias towards pigment losses. The apparent abundance of gain-of-function mutations observed in the Antirrhineae may be due to repeated introgression of adaptive alleles via rare hybridisation events with closely-related taxa, especially since homologous loci seem to control pigmentation across at multiple species (Hackbarth et al., 1942). A recent study showed that the repeated evolution of red flowers among subspecies of *Mimulus aurantiacus* was due to adaptive introgression of an allele of the transcription factor *MaMyb2* (Stankowski and Streisfeld, 2015). If a similar process occurs in Antirrhineae this would accelerate transition rates between phenotypic states, and allow for more rapid adaptive shifts to changing environments.

Our survey of the Antirrhineae revealed a marked heterogeneity in the abundances of different polymorphism types. More than half of species reported to be polymorphic for flower colour have white and anthocyanin morphs,

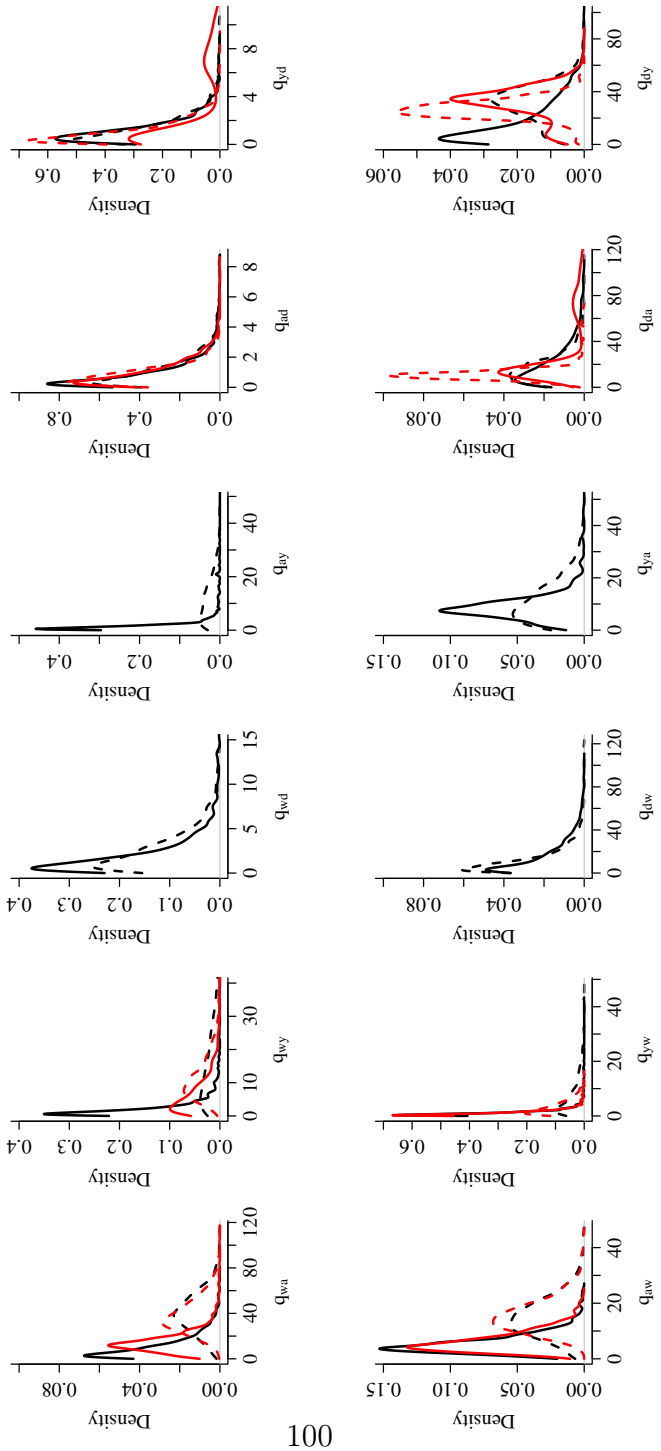
while only three taxa are have white and yellow morphs (table 5.1). This relative abundance of anthocyanin polymorphisms may be due to the much longer length of the anthocyanin pathway than the aurone pathway, and hence the greater mutational target size (Richards, 1997; Nakayama et al., 2000). A survey of the British flora show a similar pattern (Warren and Mackenzie, 2001), and our results suggest that this trend is general throughout Eurasia and North America for this group. Interestingly, this abundance of anthocyanin-white polymorphism is mirrored by consistently higher estimates of transitions between these across models and datasets (table 5.2). The abundance of allelic variants in the anthocyanin pathway may be a major driver of the high rates of transition between anthocyanin and white flowers. This suggests that the processes causing and maintaining variation in the short term may reflect evolutionary processes on longer time scales.

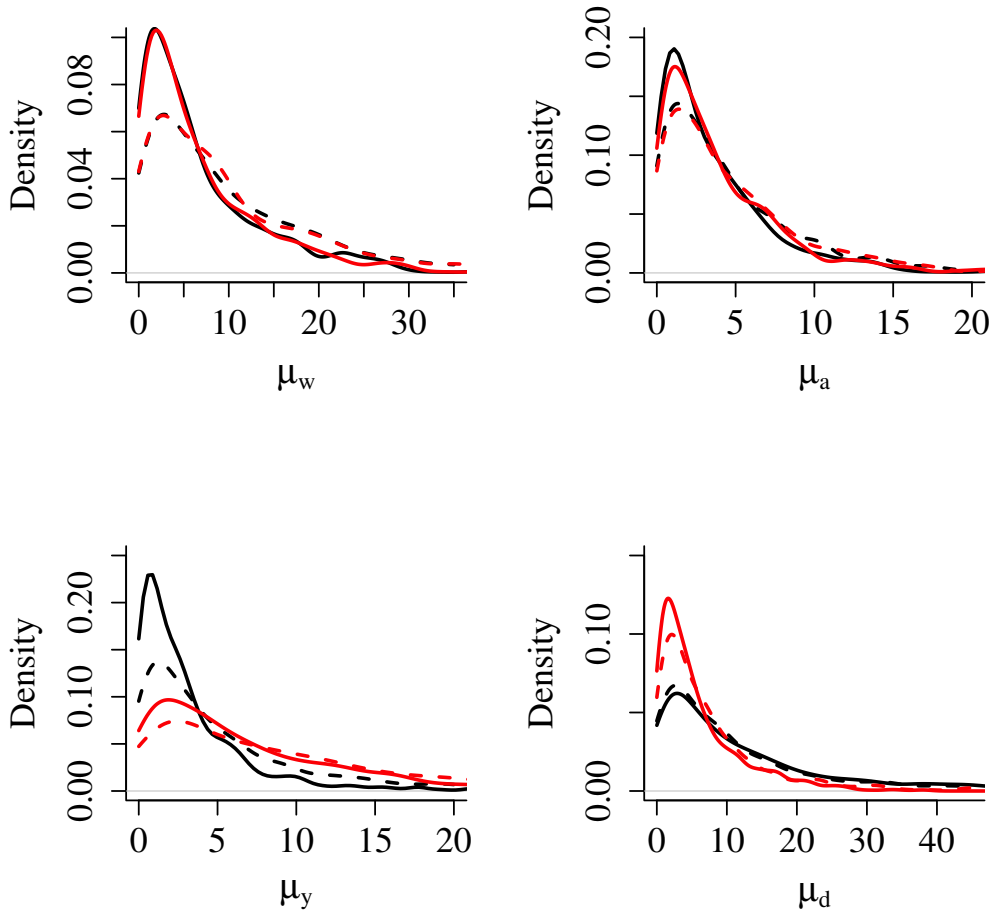
These findings highlight three areas which future efforts might focus on to improve our understanding of the evolution of floral pigmentation. Firstly, increased attention is needed to quantify the relative costs and benefits of yellow and anthocyanin pigments. This would elucidate mechanisms which might promote the spread of single-pigmented phenotypes, and whether there is some cost to double pigmentation which prevents its establishment. Secondly, we need a better understanding of the genetic architecture of flower colour transitions. Genomic approaches should provide the opportunity to sequence flower-colour genes for many species and populations. This will enable us to determine whether multiple independent alleles at these loci segregate within species or genera, and whether these alleles are able to introgress and cause multiple shifts in flower colour. Finally, the correlation between polymorphism and transition rates reflects the gradual rather than binary nature of evolutionary transitions, and highlights the need for phylogenetic comparative methods to explicitly incorporate information on phenotype or allele frequencies. Novel insight into flower-colour evolution is likely to come from considering pigmentation biology at the molecular, organismal and population levels.

Supplementary figures

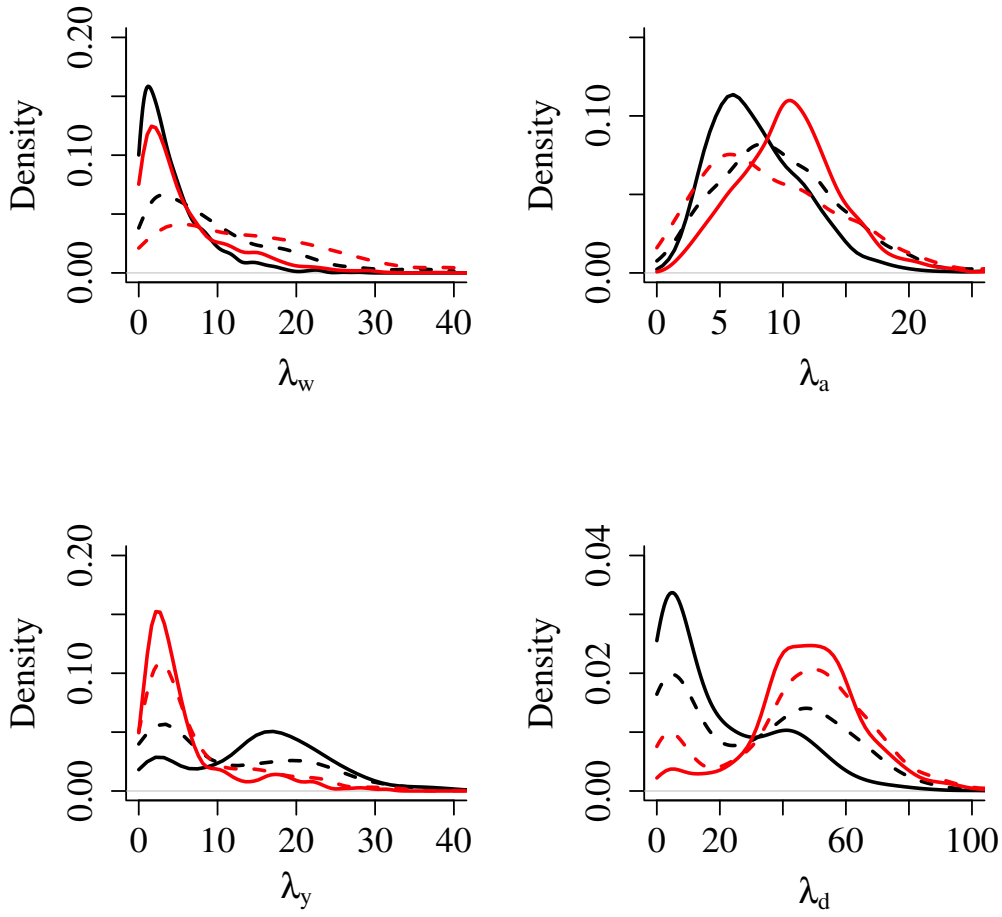
The following figures show the posterior distribution of parameter estimates from MCMC runs in diversitree. Black lines show the distribution under the full model, and red lines show distributions under the constrained model. Solid lines show distributions for monomorphic taxa only, whilst dashed lines are for datasets including polymorphic taxa. Note the different scales on both the x and y axes.

Supplementary figure 1: Posterior probability distributions for transition rate estimates.





Supplementary figure 2: Posterior probability distributions for extinction rates.



Supplementary figure 3: Posterior probability distributions for speciation rates.

Chapter 6

General discussion

6.1 Summary of main results

This thesis has examined selection on floral colour at multiple levels in a hybrid zone of the snapdragon *Antirrhinum majus*. In chapter 2 I demonstrated that bumblebee pollinators forage non-randomly on *Antirrhinum* with respect to flower colour. There was no evidence for an overall preference for particular genotypes, but plants with locally common phenotypes were visited more often than locally rare phenotypes. This matches a previous study which found that phenotypic frequency predicts the female component of plant fitness (Tastard et al., 2012), and identifies pollinators as the agent of selection. Cage experiments using näive bumblebees showed that flower colour alone can cause pollinators to forage non-randomly on snapdragons. Pollinators in experimental arrays showed floral constancy to either yellow or magenta flowers, and this constancy was sufficient to explain the degree of assortative mating among plants, measured as a deficit of hybrids in the progeny.

Using paternity analysis I found direct evidence that non-random foraging by pollinators influences the mating patterns among plants in the hybrid zone (chapter 4). The results reveal clear patterns of assortative mating for two loci controlling yellow and magenta pigmentation in the flower. Mating between two individuals was more likely the more pigmentation alleles they shared. I found no evidence for a relationship between pollen export and phenotypic frequency, suggesting that selection for phenotypic frequency acts primarily on the female, rather than male, component of reproductive suc-

cess. In order to carry out this analysis it was necessary to extend previous methods of genealogy reconstruction to allow for shared paternity within a half-sibling family, and to work efficiently with SNP genotype data (chapter 3). This allowed me to assess mating patterns in the hybrid zone whilst integrating out uncertainty regarding sibship structure and paternity.

Finally, I used a phylogenetic comparative approach to examine the patterns of floral evolution at the much broader taxonomic scale. Pigment evolution in the tribe Antirrhineae is characterised by repeated transitions between yellow and anthocyanin pigmentation. The genotypes which are rare in the the hybrid zone (those expressing either both or neither yellow and magenta pigments) are also those which are least represented across the tribe. The results point to repeated bouts of selection for flowers with a single pigment, although given the diverse roles of pigments in plants biology (Grotewold, 2006; Strauss and Whittall, 2006; Harder and Johnson, 2009), the agents of selection are likely to vary from case to case.

6.2 Maintenance of the hybrid zone

Changes in allele frequency across a hybrid zone reflect a balance between dispersal (σ) of parental populations, and the total selection (s) against immigrant alleles in a form proportional to $\sigma/\sqrt{s}$ (Haldane, 1948; Slatkin, 1973). Knowledge about dispersal can be used to directly infer the strength of total selection to maintain clines. This in turn can be compared to estimates of selection for particular traits to assess their contribution to total selection. For clines maintained by frequency-dependent selection, the selection needed to maintain a cline of width w is $(8\sigma^2)/w^2$ (Mallet and Barton, 1989a). Since pollen disperses paternal genes only, selection on pollen dispersal is half this amount. Given the one-dimensional estimate of East-West pollen dispersal distance of 645m ($\pm$40m) from chapter 4, and an estimate of cline width for *Rosea* of 2000m (David Field, unpublished data), this corresponds to a selection coefficient of 0.42 ($\pm$0.05). The estimated selection coefficient of pollinator visitation of 0.90 suggests that positive frequency-dependent selection by pollinators should be ample to maintain the clines at pigmentation loci.

Three caveats should be noted about this result. Firstly, as discussed in chapter 4, the estimate of dispersal is likely to be an overestimate. This estimate will be improved with the addition of more genotype data, and it will

be worthwhile to corroborate this estimate with other estimates of dispersal, such as inference from the decay in linkage disequilibrium across the hybrid zone (Barton and Gale, 1993). Secondly, if selection really is as strong as this, then the diffusion approximation of migration-selection balance proportional to $\sigma/\sqrt{s}$ breaks down, and will not provide reliable estimates of selection (Haldane, 1948; Slatkin, 1973). Finally, there is evidence assortative mating demonstrated in chapters 2 and 4 is reflected in widespread heterozygote deficit in the hybrid zone (David Field, unpublished data), which will sharpen the cline and reduce the estimate of selection due to positive frequency dependence. Further investigation is therefore needed into both the degree of heterozygote deficit and precise dispersal patterns to infer the strength of the positive frequency-dependent selection observed.

Frequency seems to be more important for female than male fitness. Unfortunately, Tastard et al. (2012) did not report selection coefficients for female fitness, so it is difficult to draw direct comparison with visitation. Moreover, this estimate of pollen dispersal may be an overestimate. A more robust estimate in this context would be to infer effective dispersal from linkage disequilibrium, which is less sensitive to missing long distance migration events, or variation between years (Barton and Gale, 1993). Therefore, although we confirmed that pollinators exert frequency-dependent selection on *Antirrhinum majus*, we cannot yet rule out other forms of selection.

The direct effects of frequency-dependence and assortative mating will influence a hybrid zone in two distinct ways. Selection for common morphs will drive up the frequency of dominant yellow plants in the West, and magenta plants in the East, and directly contribute to cline maintenance. Assortment will limit opportunities for gene flow, strengthening the isolation caused by clines, but without maintaining them directly. Nevertheless, as discussed in chapter 2, frequency-dependence and assortment represent a chicken-and-egg scenario. On one hand, preference purely for common morphs will lead to assortment, because pollinators move gametes between plants of the same common species. One the other hand, preference for familiar flower will promote common morphs, because pollinators are more likely to have become familiar with the most abundant floral forms encountered, meaning there is a greater pool of potential mates for common forms. Both the effects of frequency dependent foraging and floral constancy will contribute to reproductive isolation between *A. m. pseudomajus* and *A. m. striatum*.

A testable prediction of this is that pure assortment should lead to fixation of whichever form is initially most frequent. For example, *Partula* snails vary

for coil chirality. This lead directly to assortative mating because dextral and sinistral forms are physically unable to mate (Johnson, 1982). This seems to have indirectly caused locally common morphs to fix, which explains the mosaic pattern of dextral and sinistral populations observed today (Johnson, 1982). This might also explain why it is often difficult to observe frequency-dependent selection mediated by pollinators (Smithson, 2001); in most cases, one morph has already fixed. This highlights the utility of hybrid zones for investigating frequency-dependent selection, where migration from parental populations maintains variation in the face of selection.

6.3 Pollinators and reproductive isolation

This *Antirrhinum* hybrid zone contrasts with other plant systems divergent for floral traits. Most systems compare species which have diverged for a range of floral, ecological and life history traits (e.g. Fulton and Hodges, 1999; Campbell et al., 1997; Emms and Arnold, 2000; Streisfeld et al., 2013), which all influence reproductive isolation to some degree. For example, isolation between *Mimulus cardinalis* and *Mimulus lewisii* is strongly determined by their respective pollinators (hummingbirds and bumblebees), but indices of reproductive isolation for pollen and seed traits, and geographic overlap are also as strong or stronger than the indices identified in this thesis between *A. m. striatum* and *A. m. pseudomajus* (Ramsey et al., 2003; Sobel and Chen, 2014). In *A. majus,* flower colour appears to be the only trait which differs. There is no obvious difference in floral shape or flowering phenology, and the hybrid zone is perpendicular to altitude, the main axis of environmental variation. Suchet et al. (2011) found evidence of differences in floral scent between *A. m. striatum* and *A. m. pseudomajus,* but reanalysis by Jaworski (2015) could not confirm this result. This thesis has shown that flower colour sufficient to explain non-random foraging by pollinators, which in turn causes frequency-dependent selection and assortative mating among snapdragons. This represents the first stage in the evolution of reproductive isolating barriers; by reducing the homogenising effect of gene flow, this would facilitate the subsequent evolution of further barriers in the future.

Non-random foraging has the seemingly counter-intuitive implication that pollinators focus on some plant species and bypass others, even when these offer nectar and pollen rewards. There is enormous diversity in floral architecture, and pollinators need to first invest time in learning how to access

rewards, and then recall this information on subsequent foraging bouts. This places a constraint on how much information can be efficiently processed, and there may be a net gain in reward by concentrating foraging on a few plants (Waser, 1986; Chittka et al., 1999). *Antirrhinum* flowers are large and difficult to open, but produce abundant nectar (Tastard, 2009). Only large bees have the strength and intelligence to access the flowers (Vargas et al., 2010). Once they do however, they are rewarded with a rich food source which is out of reach of most competitors, and are in turn likely to visit other *Antirrhinum* plants, taking pollen with them. In experimental arrays I frequently observed pollinators to switch flower colour, only to reject the flower before opening it, suggesting that they recognise yellow and magenta flowers as being distinct. The combination of floral complexity and frequency would therefore explain why pollinators should forage non-randomly on *Antirrhinum.*

6.4 Frequency-dependent selection

If pollinators restrict foraging to a subset of available plant species, there is an expectation that they should favour common species, since these provide most total reward (Smithson and Macnair, 1996). Analogous to pollinators, predators also focus on a subset of the available prey items if this maximises net efficiency, causing frequency-dependent selection on prey (Clarke, 1962). Frequency-dependent selection by predators has been much more throughly investigated that caused by pollinators; studies in the wild are rare, and the results mixed (for reviews see Allen and Greenwood, 1988 for predation and Smithson, 2001 for pollination). In this respect, selection on *Antirrhinum* is more similar to predator-prey systems in animals than to other plant systems, such as *Mimulus, Aquilegia* or *Iris.* In the former case, fitness is determined to a large degree by the cognitive biology of the interacting animal and heavily dependent on local conditions. In the latter, fitness is more dependent on aspects of plant biology, such as floral colour and shape, and postzygotic genetic interactions (e.g. Ramsey et al., 2003). This highlights the extent to which parallels between predator-prey dynamics and pollination biology have been under-explored.

The parallel with predation is particularly clear in the case of mimicry. In classical Müllerian mimicry, species signal their distastefulness through warning colouration (Müller, 1878). Predators generalise previous negative experiences and avoid similar patterns in the future, exerting positive frequency-

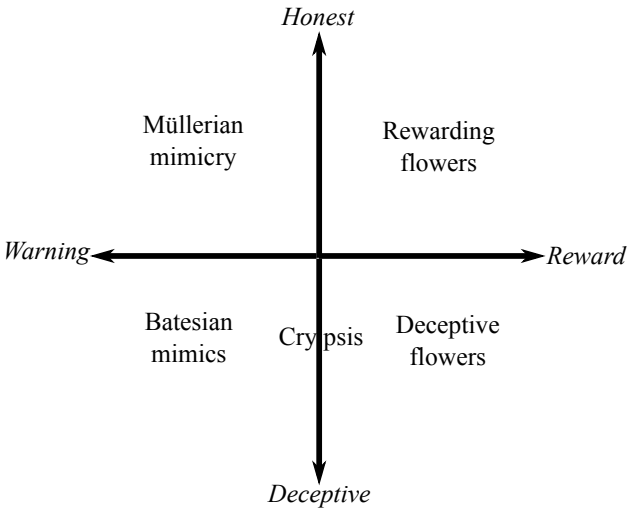


Figure 6.1: A schematic diagram illustrating signalling by rewarding flowers relative to classical mimics. The x-axis indicates the nature of the signal emitted, and the y-axis how truthful the signal is. Palatable cryptic prey are deceptive, in that they try to blend in with the environment, but make no signal of warning or reward (at least intentionally).

dependent selection on prey (Mallet and Barton, 1989b). Rewarding flowers can be viewed as a form of mimicry in that they send honest signals about nectar or pollen rewards to animals to induce them to provide pollinator service. Pollinators can rapidly learn to associate signals such as colour with reward (Heinrich, 1979), and in turn cause selection for convergent floral phenotypes across entire communities (McEwen and Vamosi, 2010a). Batesian and floral mimics, such as sexually- or food-deceptive orchids, also signal warnings or reward respectively, but send a false signal (figure 6.1; Bates, 1863; Dodson and Frymire, 1961). When deceptive mimics become common the deception is no longer effective, causing negative frequency-dependent selection (Gigord et al., 2001). The strength of frequency-dependent selection estimated for *Antirrhinum* is of a similar magnitude to estimates for *Heliconius* butterflies (s=0.52, a Müllerian mimic) and *Dactylorhiza sambucina* (s=0.59-0.66, a deceptive orchid). An advantage of investigating 'mimicry' in this sense using rewarding plants such as *Antirrhinum* is that it is easier to identify and quantify components of selection on rewarding plants, as illustrated in chapters 2 and 4; they do not run away like prey animals, they are more abundant than deceptive plants, and their floral signals make locating them straightforward.

The delineation of mimicry systems as in figure 6.1 highlights two contrasting aspects of frequency-dependent selection. Firstly, the horizontal axis distinguishes systems under selection for survival from those under selection for reproductive success. Natural and sexual selection by predators and pollinators will influence pollinators in different ways. Predators tend to exert direct frequency-dependent selection, which will alter allele frequencies in situ. In contrast, pollinators more frequently cause assortative mating among plants, with frequency-dependence a secondary outcome (Waser, 1986; Smithson, 2001). In addition to altering allele frequencies, this also influences the movement of gametes through a landscape, and reduces the pool of gametes which can unite to form a zygote. Pollinators therefore have the potential to influence both the abundance and spread of adaptive alleles.

Secondly, the vertical axis of figure 6.1 discriminates systems under positive versus negative frequency-dependent selection. Positive frequency-dependence acts to fix common alleles, reducing genetic diversity within a population. As previously noted, this predicts that evolutionary outcomes depend heavily on initial frequencies. In contrast, negative-frequency dependence acts favour rare alleles, and maintains genetic diversity within a population. The observation of positive-frequency dependence from chapter 2 is therefore at

odds with the finding of frequent transitions in flower colour across the Antirrhineae phylogeny from chapter 5. How can a new, and therefore rare, allele rise to fixation in the face of positive frequency-dependence? Selection on plant pigmentation may involve a vast range of functions beyond pollinator attraction (Strauss and Whittall, 2006), and we know almost nothing about the relative cost and benefits of different pigment classes for these. Fluctuating selection for different pigments at different times is possible, if not likely. Moreover, *Antirrhinum* does not live in a vacuum and pollinator behaviour may be influenced by the spectrum of colours in the floral community. This can either cause convergent selection for similar flower colours (McEwen and Vamosi, 2010b) or diversifying selection for distinctness (Muchhala et al., 2014) on local scales.

6.5 Further work

Although this thesis has identified pollinator behaviour as a major contributor to maintenance of the *Antirrhinum* hybrid zone, other factors may play a role. At the time of writing investigations have begun into possible differences in flowering phenology and pollen fertility across the hybrid zone, as well as possible gradients in soil quality or vegetation (David Field, Alison Kwok, Tom Ellis and Maria Claro Melo Hurtado, unpublished datasets). A potentially fruitful avenue for future work might consider the link between floral pigmentation and drought or heat stress (Warren and Mackenzie, 2001; Coberley and Rausher, 2003), both of which are likely to be important in the often dry Spanish summers. Furthermore, there are two patches on both the upper and lower roads where snapdragons are not seen at the boundary of the core hybrid zone and the yellow flank (figure 1.3). A drop in density such as this is expected to trap clines geographically and present a barrier to gene flow (Barton, 1979). This will augment the effect of positive frequency-dependent selection on either side by ensuring that immigrant alleles are at consistently at low frequency.

I was not able to investigate the behaviour of my method for sibship and paternity assignment as well as I would have liked. Further simulation work is needed to test the robustness of the model to missing fathers, the distribution of family sizes and misspecification of prior beliefs about family sizes, nor on the ultimate effect of how this would influence inference about biological parameters. It is nevertheless clear that the quality of the output of the

method depends strongly on the how much information is contained in the matrix **G**. An obvious way to improve this is to genotype more SNPs. Furthermore, there is a strong case for including other non-genetic information about paternity, especially geographic information, and to jointly estimate paternity and parameters of interest (Oddou-Muratorio et al., 2005; Hadfield et al., 2006). On one hand, this can improve inference about paternity. Given two candidate fathers, each equally likely to be the true father based on genetic information but at different distances from the mother, there is a clear prior expectation that the nearer candidate is the true fathers (Hadfield et al., 2006). At the same time, estimating paternity and then fitting a dispersal kernel to those likelihoods can cause a dispersal to be overestimated, which is alleviated when paternity and likelihood are estimated jointly (Austerlitz et al., 2004; Oddou-Muratorio et al., 2005; Hadfield et al., 2006). Joint estimation can improve inference both about pedigree structure and biological parameters of interest.

Hadfield et al. (2006) has described how to jointly estimate a pedigree and population parameters for regular parentage assignment. However, this only applies to individual offspring. Ideally, one would like to use information from shared alleles among full-siblings to aid in the identification of parents (Wang, 2007), but no method exists to do this at present. It would be straightforward to adapt the method from chapter 3 to do this by multiplying **G** by a matrix describing probabilities of paternity based on one or more biological parameters, and perform subsequent clustering. One could then use a Gibbs sampler to alter these parameters and explore the posterior distribution, much akin to Hadfield et al. (2006). The clustering technique I describe is fast, so it should be efficient to perform MCMC for the whole dataset. This would be a useful tool for future studies paternity studies in plants, where it is essential to account for family structure in half sibling families.

Bibliography

Adams, W. and D. S. Birkes, 1991. Estimating mating patterns in forest tree populations. Pp. 157–172, *in* S. Fineschi, M. Malvolti, F. Cannata, and H. Hattemer, eds. Biochemical markers in the population genetics of forest trees. SPB Academic Publishing, The Hague, The Netherlands.

Allen, J. and J. Greenwood, 1988. Frequency-dependent selection by predators [and discussion]. Philosophical Transactions of the Royal Society B 319:485–503.

Allen, J. A. and B. Clarke, 1968. Evidence for apostatic selection by wild passerines. Nature 220:501–502.

Almudevar, A. and E. C. Anderson, 2012. A new version of PRT software for sibling groups reconstruction with comments regarding several issues in the sibling reconstruction problem. Molecular Ecology Resources 12:164–178.

Almudevar, A. and C. Field, 1999. Estimation of single-generation sibling relationships based on DNA markers. Journal of Agricultural, Biological, and Environmental Statistics Pp. 136–165.

Andalo, C., M. Cruzan, C. Cazettes, B. Pujol, M. Burrus, and C. Thébaud, 2010. Post-pollination barriers do not explain the persistence of two distinct *Antirrhinum* subspecies with parapatric distributions. Plant Systematics and Evolution 286:232–234.

Anderson, E. C. and J. C. Garza, 2006. The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics 172:2567–2582.

Anderson, J., M. Wagner, C. Rushworth, K. Prasad, and T. Mitchell-Olds, 2014. The evolution of quantitative traits in complex environments. Heredity 112:4–12.

Ashman, T. and M. T. Morgan, 2004. Explaining phenotypic selection on plant attractive characters: male function, gender balance or ecological context? Proceedings of the Royal Society of London B 271:553–559.

Austerlitz, F., C. W. Dick, C. Dutech, E. K. Klein, S. Oddou-Muratorio, P. E. Smouse, and V. L. Sork, 2004. Using genetic markers to estimate the pollen dispersal curve. Molecular Ecology 13:937–954.

Barton, N. H., 1979. The dynamics of hybrid zones. Heredity 43:341–359.

Barton, N. H. and K. S. Gale, 1993. Genetic Analysis of Hybrid Zones, chap. 2, Pp. 13–45. Oxford University Press, New York.

Barton, N. H. and G. Hewitt, 1985. Analysis of hybrid zones. Annual review of Ecology and Systematics 16:113–148.

Bateman, A., 1948. Intra-sexual selection in *Drosophila*. Heredity 2:349.

———, 1951. The taxonomic discrimination of bees. Heredity 5:271–278.

Bates, D., M. Maechler, and B. Bolker, 2011. lme4: Linear mixed-effects models using S4 classes. URL `http://CRAN.R-project.org/package=lme4`.

Bates, H. W., 1863. The Naturalist on the River Amazon. Murray, London.

Baur, E., 1911. Vererbungs-und Bastardierungsversuche mit *Antirrhinum*. Zeitschrift für induktive Abstammungs- und Vererbungslehre 6:201–216.

———, 1924. Untersuchungen über das Wesen, die Entstehung und die Vererbung von Rassenunterschieden bei *Antirrhinum majus*. Gebrüder Borntträger, Leipzig.

Bell, G., 1985. On the function of flowers. Proceedings of the Royal Society of London B 224:223–265.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, B. A. Rapp, and D. L. Wheeler, 2000. Genbank. Nucleic acids research 28:15–18.

Blouin, M. S., 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology & Evolution 18:503–511.

Bowler, P. J., 1989. Evolution: the history of an idea. University of California Press.

Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen, and F. Roux, 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genetics 6:e1000940.

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

Brown, B. A. and M. T. Clegg, 1984. Influence of flower color polymorphism on genetic transmission in a natural population of the common morning glory, *Ipomoea purpurea*. Evolution 38:796–803.

Burczyk, J., W. Adams, G. Moran, and A. Griffin, 2002. Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. Molecular Ecology 11:2379–2391.

Burd, M., 1994. Bateman's principle and plant reproduction: the role of pollen limitation in fruit and seed set. The Botanical Review 60:83–139.

Burdon, J., D. Marshall, and A. Brown, 1983. Demographic and genetic changes in populations of *Echium plantagineum*. Journal of Ecology 71:667–679.

Butler, K., C. Field, C. Herbinger, and B. Smith, 2004. Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data. Molecular Ecology 13:1589–1600.

Campbell, D. R., 2004. Natural selection in *Ipomopsis* hybrid zones: implications for ecological speciation. New Phytologist 161:83–90.

Campbell, D. R., N. M. Waser, and E. J. Meléndez-Ackerman, 1997. Analyzing pollinator-mediated selection in a plant hybrid zone: hummingbird visitation patterns on three spatial scales. American Naturalist 149:295–315.

Chavannes, E. L., 1833. Monographie des Antirrhinées. Treuttel et Würtz, Paris.

Chittka, L., J. D. Thomson, and N. M. Waser, 1999. Flower constancy, insect psychology, and plant evolution. Naturwissenschaften 86:361–377.

Clark, J. S., 1998. Why trees migrate so fast: Confronting theory with dispersal biology and the paleorecord. The American Naturalist 152:204–224.

Clarke, B., 1962. Natural selection in mixed populations of two polymorphic snails. Heredity 17:319–345.

Coberley, L. and M. Rausher, 2003. Analysis of a chalcone synthase mutant in *Ipomoea purpurea* reveals a novel function for flavonoids: amelioration of heat stress. Molecular Ecology 12:1113–1114.

Coen, E. S., R. Carpenter, and C. Martin, 1986. Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. Cell 47:285–296.

Coltman, D. W., J. G. Pilkington, J. A. Smith, and J. M. Pemberton, 1999. Parasite-mediated selection against inbred Soay sheep in a free-living, island population. Evolution 53:1259–1267.

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

Coulson, T., E. A. Catchpole, S. D. Albon, B. J. Morgan, J. Pemberton, T. H. Clutton-Brock, M. Crawley, and B. Grenfell, 2001. Age, sex, density, winter weather, and population crashes in soay sheep. Science 292:1528–1531.

Coyne, J. A., H. A. Orr, et al., 2004. Speciation. Sinauer Associates Sunderland, Massachusetts.

Darwin, C., 1859. The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life. J. Murray, London.

———, 1876. The effects of cross and self fertilisation in the vegetable kingdom. J. Murray, London.

Devlin, B. and N. C. Ellstrand, 1990. Male and female fertility variation in wild radish, a hermaphrodite. American Naturalist 136:87–107.

Devlin, B., K. Roeder, and N. Ellstrand, 1988. Fractional paternity assignment: theoretical development and comparison to other methods. Theoretical and Applied Genetics 76:369–380.

Dick, C. W., 2001. Genetic rescue of remnant tropical trees by an alien pollinator. Proceedings of the Royal Society of London B 268:2391–2396.

Dittmar, E. L., C. G. Oakley, J. Ågren, and D. W. Schemske, 2014. Flowering time QTL in natural populations of *Arabidopsis thaliana* and implications for their adaptive value. Molecular ecology 23:4291–4303.

Dobzhansky, T., 1937. Genetics and the Origin of Species. Columbia University Press.

Dodson, C. and G. Frymire, 1961. Natural pollination of orchids. Missouri Botanical Garden Bulletin 49:133–152.

Eckhart, V. M., N. S. Rushing, G. M. Hart, and J. D. Hansen, 2006. Frequency–dependent pollinator foraging in polymorphic *Clarkia xantiana ssp. xantiana* populations: Implications for flower colour evolution and pollinator interactions. Oikos 112:412–421.

Elisens, W. J., 1985. Monograph of the Maurandyinae (Scrophulariaceae-Antirrhineae). Systematic Botany Monographs 5:1–97.

Emery, A., I. Wilson, S. Craig, P. Boyle, and L. Noble, 2001. Assignment of paternity groups without access to parental genotypes: Multiple mating and developmental plasticity in squid. Molecular Ecology 10:1265–1278.

Emms, S. and M. Arnold, 1997. The effect of habitat on parental and hybrid fitness: transplant experiments with Louisiana irises. Evolution 51:1112–1119.

Emms, S. K. and M. L. Arnold, 2000. Site-to-site differences in pollinator visitation patterns in a Louisiana iris hybrid zone. Oikos 91:568–578.

Epperson, B. K. and M. T. Clegg, 1987. Frequency-dependent variation for outcrossing rate among flower-color morphs of *Ipomoea purpurea*. Evolution 41:1302–1311.

Faegri, K. and L. Van der Pijl, 1966. The principles of pollination ecology. Pergamon Press, Oxford.

Fernandéz Casos, J., 1988. Asientos para una flora occidental. Fontqueria 20:63.

Fernández-Mazuecos, M., J. L. Blanco-Pastor, and P. Vargas, 2013. A phylogeny of toadflaxes (*Linaria* Mill.) based on nuclear internal transcribed spacer sequences: Systematic and evolutionary consequences. International Journal of Plant Scence. 174:234–249.

Fisher, R. A., 1930. The genetical theory of natural selection. Oxford University Press.

FitzJohn, R. G., 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods in Ecology and Evolution 3:1084–1092.

FitzJohn, R. G., W. P. Maddison, and S. P. Otto, 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Systematic Biology 58:595–611.

Friel, N. and J. Wyse, 2012. Estimating the evidence: a review. Statistica Neerlandica 66:288–308.

Fulton, M. and S. A. Hodges, 1999. Floral isolation between *Aquilegia formosa* and *Aquilegia pubescens*. Proceedings of the Royal Society of London B 266:2247–2252.

Geissman, T., E. C. Jorgensen, and B. L. Johnson, 1954. The chemistry of flower pigmentation in *Antirrhinum majus* color genotypes. I. The flavonoid components of the homozygous P, M, Y color types. Archives of biochemistry and biophysics 49:368–388.

Gigord, L. D., M. R. Macnair, and A. Smithson, 2001. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid emphDactylorhiza sambucina (L.) Soo. Proceedings of the National Academy of Sciences 98:6253–6255.

Goldberg, E. E. and B. Igić, 2008. On phylogenetic tests of irreversible evolution. Evolution 62:2727–2741.

Grant, P. R. and B. R. Grant, 2002. Unpredictable evolution in a 30-year study of darwin's finches. Science 296:707–711.

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

———, 1950. The flower constancy of bees. The Botanical Review 16:379–398.

Green, A. A., J. R. Kennaway, A. I. Hanna, J. A. Bangham, and E. Coen, 2010. Genetic control of organ shape and tissue polarity. PLoS Biology 8:e1000537.

Grotewold, E., 2006. The genetics and biochemistry of floral pigments. Annual Review of Plant Biology 57:761–780.

Güemes, J., 1994. *Antirrhinum subbaeticum* Güemes, Mateu & Sánchez-Gómez (Scrophulariaceae), especie nueva de la península Ibérica. Anales del Jardin Botanico de Madrid 51:237–247.

Hackbarth, J., P. Michaelis, and G. Scheller, 1942. Untersuchungen an dem *Antirrhinum*-Wildsippensortiment von E. Baur. Zeitschrif für induktive Abstammung und Vererbungslehre Pp. 1–102.

Hadfield, J., D. Richardson, and T. Burke, 2006. Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. Molecular Ecology 15:3715–3730.

Hagen, M., M. Wikelski, and W. D. Kissling, 2011. Space use of bumblebees (*Bombus spp.*) revealed by radio-tracking. PLoS ONE 6:e19997.

Haldane, J., 1948. The theory of a cline. Journal of genetics 48:277–284.

Harborne, J., 1966. Comparative biochemistry of flavonoids I.: Distribution of chalcone and aurone pigments in plants. Phytochemistry 5:111–115.

Harder, L. D. and S. D. Johnson, 2009. Darwin's beautiful contrivances: evolutionary and functional evidence for floral adaptation. New Phytologist 183:530–545.

Hart, M. W., 2011. The species concept as an emergent property of population biology. Evolution 65:613–616.

Heinrich, B., 1979. "Majoring" and" minoring" by foraging bumblebees, *Bombus vagans*: an experimental analysis. Ecology 60:246–255.

Hewitt, G. M., 1988. Hybrid zones: Natural laboratories for evolutionary studies. Trends in Ecology & Evolution 3:158–167.

Hodges, S. A. and M. L. Arnold, 1994. Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens*. Proceedings of the National Academy of Sciences 91:2493–2496.

Hodgins, K. A. and S. C. Barrett, 2008. Natural selection on floral traits through male and female function in wild populations of the heterostylous daffodil *Narcissus triandrus*. Evolution 62:1751–1763.

Hopkins, R. and M. D. Rausher, 2012. Pollinator-mediated selection on flower color allele drives reinforcement. Science 335:1090–1092.

Irwin, R. E., S. Y. Strauss, S. Storz, A. Emerson, and G. Guibert, 2003. The role of herbivores in the maintenance of a flower color polymorphism in wild radish. Ecology 84:1733–1743.

Jaworski, C., 2015. Interactions plantes-insectes dans le réseau d'espèces pyrénéen d'Antirrhinum majus: métacommunauté, comportement et odeurs florales. Ph.D. thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier.

Jaworski, C. C., C. Andalo, C. Raynaud, V. Simon, C. Thébaud, and J. Chave, 2015. The influence of prior learning experience on pollinator choice: An experiment using bumblebees on two wild floral types of *Antirrhinum majus*. PloS ONE 10:e0130225.

Johnson, M. S., 1982. Polymorphism for direction of coil in *Partula suturalis*: behavioural isolation and positive frequency dependent selection. Heredity 49:145–151.

Jones, A. G., C. M. Small, K. A. Paczolt, and N. L. Ratterman, 2010. A practical guide to methods of parentage analysis. Molecular ecology resources 10:6–30.

Jones, K. N., 1996a. Fertility selection on a discrete floral polymorphism in *Clarkia* (Onagraceae). Evolution 50:71–79.

———, 1996b. Pollinator behavior and postpollination reproductive success in alternative floral phenotypes of *Clarkia gracilis* (Onagraceae). International Journal of Plant Sciences 157:733–738.

Joron, M., C. Jiggins, A. Papanicolaou, and W. McMillan, 2006. *Heliconius* wing patterns: an evo-devo model for understanding phenotypic diversity. Heredity 97:157–167.

Katoh, K. and D. M. Standley, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780.

Key, K., 1968. The concept of stasipatric speciation. Systematic Biology 17:14–22.

Kobayashi, S., K. Inoue, and M. Kato, 1999. Mechanism of selection favoring a wide tubular corolla in *Campaula punctata*. Evolution 53:752–757.

Kruschke, J., 2010. Doing Bayesian data analysis: a tutorial introduction with R. Academic Press, Oxford.

Lande, R. and S. J. Arnold, 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lepais, O., B. Darvill, S. O'Connor, J. L. Osborne, R. A. Sanderson, J. Cussans, L. Goffe, and D. Goulson, 2010. Estimation of bumblebee queen dispersal distances using sibship reconstruction method. Molecular Ecology 19:819–831.

Levin, D. A., 1972. Low frequency disadvantage in the exploitation of pollinators by corolla variants in *Phlox*. American Naturalist 106:453–460.

Levin, D. A. and H. W. Kerster, 1970. Phenotypic dimorphism and populational fitness in *Phlox*. Evolution Pp. 128–134.

———, 1974. Gene flow in seed plants. Pp. 139–220, *in* T. Dobzhansky, M. Hecht, and W. Steere, eds. Evolutionary Biology. Plenum Press.

Maddison, W. P., P. E. Midford, and S. P. Otto, 2007. Estimating a binary character's effect on speciation and extinction. Systematic Biology 56:701–710.

Mallet, J. and N. Barton, 1989a. Inference from clines stabilized by frequency-dependent selection. Genetics 122:967–976.

Mallet, J. and N. H. Barton, 1989b. Strong natural selection in a warning-color hybrid zone. Evolution 24:421–431.

Marshall, T., J. Slate, L. Kruuk, and J. Pemberton, 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7:639–655.

Mather, K., 1947. Species crosses in *Antirrhinum* I. Genetic isolation of the species *majus, glutinosum* and *orontium*. Heredity 1:175–186.

Mayr, E., 1942. Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press.

McEwen, J. R. and J. C. Vamosi, 2010a. Floral colour versus phylogeny in structuring subalpine flowering communities. Proceedings of the Royal Society B 277:2957–2965.

———, 2010b. Floral colour versus phylogeny in structuring subalpine flowering communities. Proceedings of the Royal Society B 277:2957–2965.

Meagher, T. R., 1986. Analysis of paternity within a natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents. American Naturalist 128:199.

Meagher, T. R. and E. Thompson, 1986. The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theoretical Population Biology 29:87–106.

Meléndez-Ackerman, E. and D. R. Campbell, 1998. Adaptive significance of flower color and inter-trait correlations in an *Ipomopsis* hybrid zone. Evolution 52:1293–1303.

Mitchell-Olds, T. and R. G. Shaw, 1987. Regression analysis of natural selection: statistical inference and biological interpretation. Evolution 41:1149–1161.

Muchhala, N., S. Johnsen, and S. D. Smith, 2014. Competition for hummingbird pollination shapes flower color variation in Andean Solanaceae. Evolution 68:2275–2286.

Müller, F., 1878. Über die Vortheile der Mimikry bei Smetterlingen. Zoologischer Anzeiger 1:54–55.

Munz, P. A., 1926. The Antirrhinoideae-Antirrhineae of the New World. Proceedings of the Californian Academy of Science 15:323–397.

Nakayama, T., 2002. Enzymology of aurone biosynthesis. Journal of bioscience and bioengineering 94:487–491.

Nakayama, T., K. Yonekura-Sakakibara, T. Sato, S. Kikuchi, Y. Fukui, M. Fukuchi-Mizutani, T. Ueda, M. Nakao, Y. Tanaka, T. Kusumi, et al., 2000. Aureusidin synthase: a polyphenol oxidase homolog responsible for flower coloration. Science 290:1163–1166.

Newton, M. A. and A. E. Raftery, 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. Journal of the Royal Statistical Society B 56:3–48.

Nielsen, R., D. K. Mattila, P. J. Clapham, and P. J. Palsbøll, 2001. Statistical approaches to paternity analysis in natural populations and applications to the north atlantic humpback whale. Genetics 157:1673–1682.

Oddou-Muratorio, S., E. K. Klein, and F. Austerlitz, 2005. Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. Molecular Ecology 14:4441–4452.

Ohmiya, A., S. Kishimoto, R. Aida, S. Yoshioka, and K. Sumitomo, 2006. Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in *Chrysanthemum petals*. Plant Physiology 142:1193–1201.

Ono, E., M. Fukuchi-Mizutani, N. Nakamura, Y. Fukui, K. Yonekura-Sakakibara, M. Yamaguchi, T. Nakayama, T. Tanaka, T. Kusumi, and Y. Tanaka, 2006. Yellow flowers generated by expression of the aurone biosynthetic pathway. Proceedings of the National Academy of Sciences 103:11075–11080.

Onslow, M. W., 1916. The anthocyanin pigments of plants. Cambridge university press.

Orr, H. A., 1996. Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335.

Osborne, J. L., A. P. Martin, N. L. Carreck, J. L. Swain, M. E. Knight, D. Goulson, R. J. Hale, and R. A. Sanderson, 2008. Bumblebee flight distances in relation to the forage landscape. Journal of Animal Ecology 77:406–415.

Pagel, M., 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proceedings of the Royal Society of London B 255:37–45.

Painter, I., 1997. Sibship reconstruction without parental information. Journal of Agricultural, Biological, and Environmental Statistics Pp. 212–229.

Pemberton, J., 2008. Wild pedigrees: the way forward. Proceedings of the Royal Society of London B 275:613–621.

Pennell, F. W., 1947. Some hitherto undescribed Scrophulariaceae of the Pacific states. Proceedings of the Academy of Natural Sciences of Philadelphia 99:155–199.

Politis, D. N. and J. P. Romano, 1994. Large sample confidence regions based on subsamples under minimal assumptions. The Annals of Statistics 22:2031–2050.

Proctor, M. and P. Yeo, 1973. The pollination of flowers. Collins, London.

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

R Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org`.

Rabosky, D. L. and E. E. Goldberg, 2015. Model Inadequacy and Mistaken Inferences of Trait-Dependent Speciation. Systematic Biology 64:340–355.

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

Rausher, M. D., 2006. The evolution of flavonoids and their genes. Pp. 175–211, *in* E. Grotewald, ed. The Science of Flavonoids. Springer, London.

———, 2008. Evolutionary transitions in floral color. Internation Journal of Plant Science 169:7–21.

Richards, A. J., 1997. Plant breeding systems. Garland Science.

Richmond, M. L., 2001. Women in the early history of genetics: William Bateson and the Newnham College Mendelians, 1900-1910. Isis Pp. 55–90.

Ronquist, F. and J. P. Huelsenbeck, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rothmaler, W., 1956. Taxonomische Monographie der Gattung *Antirrhinum*. Akademie Verlag, Berlin.

Schemske, D. W. and P. Bierzychudek, 2001. Evolution of flower color in the desert annual *Linanthus parryae*: Wright revisited. Evolution 55:1269–1282.

Schoen, D. J. and S. C. Stewart, 1986. Variation in male reproductive investment and male reproductive success in white spruce. Evolution 40:1109–1120.

Schwarz-Sommer, Z., B. Davies, and A. Hudson, 2003. An everlasting pioneer: the story of *Antirrhinum* research. Nature Reviews Genetics 4:657–666.

Schwinn, K., J. Venail, Y. Shang, S. Mackay, V. Alm, E. Butelli, R. Oyama, P. Bailey, K. Davies, and C. Martin, 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. Plant Cell 18:831–851.

Slatkin, M., 1973. Gene flow and selection in a cline. Genetics 75:733–756.

Smith, B. R., C. M. Herbinger, and H. R. Merry, 2001. Accurate partition of individuals into full-sib families from genetic data without parental information. Genetics 158:1329–1338.

Smith, S. D. and E. E. Goldberg, 2015. Tempo and mode of flower color evolution. American Journal of Botany 102:1014–1025.

Smith, S. D., R. E. Miller, S. P. Otto, R. G. FitzJohn, and M. D. Rausher, 2010. The effects of flower color transitions on diversification rates in morning glories (*Ipomoea* subg. Quamoclit, Convolvulaceae). Pp. 202–226, *in* M. Long, H. Gu, and Z. Zhou, eds. Darwin's Heritage Today. Higher Education Press, Beijing, China.

Smith, S. D. and M. D. Rausher, 2011. Gene loss and parallel evolution contribute to species difference in flower color. Molecular Biology and Evolution 28:2799–2810.

Smithson, A., 2001. Pollinator preference, frequency dependence, and floral evolution. chap. 12, Pp. 237–258, *in* L. Chittka and J. D. Thompson, eds. Cognitive Ecology of Behvaiour: Animal Behaviour and Floral Evolution. Cambridge University Press, New York.

Smithson, A. and M. Macnair, 1996. Frequency-dependent selection by pollinators: mechanisms and consequences with regard to behaviour of bumblebees *Bombus terrestris*(L.)(Hymenoptera: Apidae). Journal of Evolutionary Biology 9:571–588.

———, 1997. Density-dependent and frequency-dependent selection by bumblebees *Bombus terrestris*(L.)(Hymenoptera: Apidae). Biological Journal of the Linnean Society 60:401–417.

Smouse, P., T. Meagher, and C. Kobak, 1999. Parentage analysis in *Chamaelirium luteum* (L.) Gray (Liliaceae): why do some males have higher reproductive contributions? Journal of Evolutionary Biology 12:1069–1077.

Sobel, J. M. and G. F. Chen, 2014. Unification of methods for estimating the strength of reproductive isolation. Evolution 68:1511–1522.

Sokal, R. R. and C. Michener, 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38:1409–1438.

Spaethe, J., J. Tautz, and L. Chittka, 2001. Visual constraints in foraging bumblebees: flower size and color affect search time and flight behavior. Proceedings of the National Academy of Sciences 98:3898–3903.

Speta, F., 1980. Die Gattungen *Chaenorhinum (DC.)* REICHENB. und *Micorrhinum* (ENDL.) FOURR. im östlichen Teil ihrer Areale (Balkan bis Indien). Stapfia Pp. 1–72.

Stankowski, S. and M. A. Streisfeld, 2015. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. Proceedings of the Royal Society of London. B 282:20151666.

Stanton, M. L., 1987a. Reproductive biology of petal color variants in wild populations of *Raphanus sativus*: I. Pollinator response to color morphs. American Journal of Botany 74:178–187.

———, 1987b. Reproductive biology of petal color variants in wild populations of *Raphanus sativus*: I. Pollinator response to color morphs. American Journal of Botany 74:178–187.

Stanton, M. L., A. A. Snow, and S. N. Handel, 1986. Floral evolution: attractiveness to pollinators increases male fitness. Science 232:1625–1627.

Stanton, M. L., A. A. Snow, S. N. Handel, and J. Bereczky, 1989. The impact of a flower-color polymorphism on mating patterns in experimental populations of wild radish (*Raphanus raphanistrum* L.). Evolution 43:335–346.

Stinchcombe, J. R., 2014. Cross-pollination of plants and animals: wild quantitative genetics and plant evolutionary genetics. Pp. 128–146, *in* A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative Genetics in the Wild. Oxford University Press.

Strauss, S. Y., R. E. Irwin, and V. M. Lambrix, 2004. Optimal defence theory and flower petal colour predict variation in the secondary chemistry of wild radish. Journal of Ecology 92:132–141.

Strauss, S. Y. and J. B. Whittall, 2006. Non-pollinator agents of selection on floral traits. Pp. 120–138, *in* L. Harder and S. Barrett, eds. Ecology and Evolution of Flowers. Oxford University Press, Oxford.

Streisfeld, M. A. and M. D. Rausher, 2011. Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. Evolution 65:629–642.

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

Stubbe, H., 1966. Genetik und Zytologie von *Antirrhinum* L. sect. *Antrirrhinum*. Gustav Fischer Verlag, Jena.

Suchet, C., L. Dormont, B. Schatz, M. Giurfa, V. Simon, C. Raynaud, and J. Chave, 2011. Floral scent variation in two *Antirrhinum majus* subspecies influences the choice of naïve bumblebees. Behavioral Ecology and Sociobiology 65:1015–1027.

Sutton, D. A., 1988. A revision of the tribe Antirrhineae. British Museum (Natural History).

Szymura, J. M. and N. H. Barton, 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. Evolution 45:237–261.

Tanaka, Y., S. Tsuda, and T. Kusumi, 1998. Metabolic engineering to modify flower color. Plant Cellular Physiology 39:1119–1126.

Tastard, E., 2009. Maintien d'une zone hybride de gueules de loup *(Antirrhinum majus)*: rôle de quelques interactions biologiques. Ph.D. thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier.

Tastard, E., C. Andalo, M. Giurfa, M. Burrus, and C. Thébaud, 2008. Flower colour variation across a hybrid zone in *Antirrhinum* as perceived by bumblebee pollinators. Arthropod-Plant Interactions 2:237–246.

Tastard, E., J. Ferdy, M. Burrus, C. Thébaud, and C. Andalo, 2012. Patterns of floral colour neighbourhood and their effects on female reproductive success in an *Antirrhinum* hybrid zone. Journal of Evolutionary Biology 25:388–399.

Tavares, H., 2014. Evolutionary genetics and genomics of flower colour loci in an Antirrhinum hybrid zone. Ph.D. thesis, John Innes Centre, University of East Anglia.

Thomas, S. C. and W. G. Hill, 2000. Estimating quantitative genetic parameters using sibships reconstructed from marker data. Genetics 155:1961–1972.

———, 2002. Sibship reconstruction in hierarchical population structures using Markov chain Monte Carlo techniques. Genetical research 79:227–234.

Thompson, D., 1988. Systematics of *Antirrhinum* (Scrophulariaceae) in the New World. Systematic Botany Monographs 22:1–142.

Thompson, E., 1975. The estimation of pairwise relationships. Annals of human genetics 39:173–188.

———, 1976a. A paradox of genealogical inference. Advances in Applied Probability 8:648–650.

Thompson, E. A., 1976b. Inference of genealogical structure. Social Science Information 15:477–526.

Tjebbes, K., 1929. Species crosses in the genus *Linaria*. Beretning on det 18. skandinav. Naturforskermøde Pp. 529–534.

Valdés, B., 1970. Flavonoid pigments in flower and leaf of the genus *Linaria*. Phytochemistry 9:1253 – 1260.

Vargas, P., C. Ornosa, F. Ortiz-Sanchez, and J. Arroyo, 2010. Is the occluded corolla of *Antirrhinum* bee-specialized? Journal of Natural History 44:1427–1443.

Wang, J., 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166:1963–1979.

———, 2007. Parentage and sibship exclusions: higher statistical power with more family members. Heredity 99:205–217.

———, 2012. Computationally efficient sibship and parentage assignment from multilocus marker data. Genetics 191:183–194.

Wang, J. and A. W. Santure, 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181:1579–1594.

Warren, J. and S. Mackenzie, 2001. Why are all colour combinations not equally represented as flower-colour polymorphisms? New Phytologist 151:237–241.

Waser, N. M., 1986. Flower constancy: definition, cause, and measurement. American Naturalist 127:593–603.

Waser, N. M. and M. V. Price, 1981. Pollinator choice and stabilizing selection for flower color in *Delphinium nelsonii*. Evolution 32:376–390.

———, 1983. Pollinator behaviour and natural selection for flower colour in *Delphinium nelsonii*. Nature 302:422–424.

Weinig, C., M. C. Ungerer, L. A. Dorn, N. C. Kane, Y. Toyonaga, S. S. Halldorsdottir, T. F. Mackay, M. D. Purugganan, and J. Schmitt, 2002. Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. Genetics 162:1875–1884.

Wessinger, C. A. and M. D. Rausher, 2012. Lessons from flower colour evolution on targets of selection. Journal of Experimental Botany 63:5741–5749.

———, 2013. Predictability and irreversibility of genetic changes associated with flower color evolution in *Penstemon bartbatus*. Evolution 68:1058–1070.

Wheldale, M., 1907. The inheritance of flower colour in *Antirrhinum majus*. Proceedings of the Royal Society of London B 79:288–305.

Whibley, A. C., 2004. Molecular and genetic variation underlying the evolution of flower colour in Antirrhinum. Ph.D. thesis, University of East Anglia.

Whibley, A. C., N. B. Langlade, C. Andalo, A. I. Hanna, A. Bangham, C. Thébaud, and E. Coen, 2006. Evolutionary paths underlying flower color variation in *Antirrhinum*. Science 313:963–966.

Whittall, J. B., C. Voelckel, D. J. Kliebenstein, and S. A. Hodges, 2006. Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. Molecular Ecology 15:4645–4657.

Wienand, U., H. Sommer, Z. Schwarz, N. Shepherd, H. Saedler, F. Kreuzaler, H. Ragg, E. Fautz, K. Hahlbrock, B. Harrison, et al., 1982. A general method to identify plant structural genes among genomic DNA clones using transposable element induced mutations. Molecular and General Genetics MGG 187:195–201.

Wright, J. and T. Meagher, 2004. Selection on floral characters in natural Spanish populations of *Silene latifolia*. Journal of Evolutionary Biology 17:382–395.

Wright, S., 1943. Isolation by distance. Genetics 28:114.

Xue, Y., R. Carpenter, H. G. Dickinson, and E. S. Coen, 1996. Origin of allelic diversity in *Antirrhinum* S locus RNases. The Plant Cell 8:805–814.

Zhang, B., C. Liu, Y. Wang, X. Yao, F. Wang, J. Wu, G. J. King, and K. Liu, 2015. Disruption of a CAROTENOID CLEAVAGE DIOXYGENASE 4 gene converts flower colour from white to yellow in *Brassica species*. New Phytologist 206:1513–1526.

Zufall, R. A. and M. D. Rausher, 2004. Genetic changes associated with floral adaptation restrict future evolutionary potential. Nature 428:847–850.

Zuur, A., E. Ieno, N. Walker, A. Saveliev, and G. Smith, 2009. Mixed effects models and extensions in ecology with R. Springer.