Macroevolutionary patterns of pollination accuracy: a comparison of three genera

W. Scott Armbruster^{1,2,3}, Christophe Pélabon⁴, Thomas F. Hansen⁵ and Geir H. Bolstad⁴

¹School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DY, UK; ²Department of Biology, NTNU, N-7491, Trondheim, Norway;

³Institute of Arctic Biology, University of Alaska, Fairbanks, AK 99775, USA; ⁴Centre for Conservation Biology, Department of Biology, NTNU, N-7491,

Trondheim, Norway; ⁵Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, PO Box 1066, N-0316 Oslo, Norway

Summary

Author for correspondence:

W. Scott Armbruster

Tel: 44 (0)23 92842081

Email: Scott.Armbruster@port.ac.uk

Received: 1 February 2009

Accepted: 18 May 2009

New Phytologist (2009) **183**: 600–617
doi: 10.1111/j.1469-8137.2009.02930.x

Key words: adaptive accuracy, adaptive surface, *Collinsia*, *Dalechampia*, optimality, pollination, precision, *Stylidium*

- We hypothesize that pollination efficiency selects for equal distances between the pollinator reward and the anthers, and the stigmas, creating an adaptive ridge. We predict that this fitness surface governs the divergence of many plant species. We use the theory of adaptive accuracy, precision and mean optimality to assess how close populations lie to the hypothesized adaptive ridge and which factors contribute to departure from the optimum.
- Patterns of accuracy of pollen placement and receipt were compared across species in three study systems, *Dalechampia* (Euphorbiaceae), *Collinsia* (Plantaginaceae) and *Stylidium* (Stylidiaceae), in order to assess the roles of stamen/stigma imprecision and population mean departure from the optimum in the generation of floral inaccuracy.
- We found that population mean departure from the optimum was the most important factor in *Dalechampia*, female imprecision and departure from the optimum were about equally important factors in *Collinsia*, and stamen and stigma imprecision were equally important in *Stylidium*, with virtually no departure from the optimum.
- Possible reasons for imprecision and departure from the optimum were assessed using phylogenetically informed methods, indicating important roles of limited floral integration in the generation of imprecision, and conflicting selective pressures, associated with outcrossing, in the generation of departure from the optimum.

Introduction

Although many fundamental questions in evolutionary biology remain unanswered, one of the most compelling is: what is the relative importance of adaptation, genetic constraints and historical contingency in the divergence of populations and species (Williams, 1992; Schluter, 2000; Gould, 2002)? The diversity of floral 'design' among plant species has been invoked repeatedly as one of the most dramatic examples of the diversification of species by natural selection (for example, Darwin, 1877; Stebbins, 1951, 1974), and constitutes a good study system to address challenging macroevolutionary questions. Since Darwin's time, it has been largely assumed that floral diversification among species reflects adaptive evolution and speciation in response to divergent selection exerted by pollinators (Grant, 1971; Stebbins, 1974; Schluter, 2000; Gavrilov, 2004). Some recent studies have suggested, however, that pollinators may be only one of several possible evolutionary

forces generating floral diversity (for example, Armbruster, 1991, 2002; Strauss & Irwin, 2004). Instead, the potential importance of developmental constraints and genetic factors, such as pleiotropy, as well as selection by other interactors, forces us to assess floral diversity more cautiously, with adaptive diversification being just one of several possible contributors. In the present contribution, we explore the role of adaptation as a probable factor, but in the context of being only one of several testable hypotheses for floral divergence.

The treatment of adaptive evolution as the movement of populations on an adaptive landscape (or surface) in allele frequency (Wright, 1931) or morphological space (Simpson, 1944) represents a useful way to conceptualize evolution in response to natural selection. The adaptive surface concept has been applied largely to assess components of relative fitness of particular gene frequencies or values of selected morphological traits for individuals in a population (Lande & Arnold, 1983; Schluter, 1988; Schluter & Nychka, 1994; O'Connell

& Johnston, 1998). This concept can, however, also be applied to the adaptive divergence of populations and species, under the assumption that they all experience the same pre-existing adaptive landscape for the components of fitness and traits being examined (for example, with multiple peaks or bumpy ridges; Armbruster, 1990; Arnold *et al.*, 2001). Although this is not always the case, it appears to be true for related species with similar ecologies, and is thus a useful conceptual approach to study the roles of adaptation, constraint, history and randomness in macroevolution. It is important to note that adaptive optima (high points on the landscape) can pertain conceptually to the totality of biological functions and traits, but must usually apply operationally to only subsets of functions and traits (components of fitness). In the present study, we consider the shape of the adaptive surface governing one function and two traits: pollination performance (a component of reproductive fitness) in response to the positions of anthers and stigmas in flowers.

A related, complementary approach to the assessment of an organism's position on its adaptive surface is to evaluate the 'adaptive accuracy' of populations. This requires the identification of adaptive optima (for fitness components), estimation of trait deviation from these optima and the assessment of the possible causes of maladaptation. The latter involves the decomposition of adaptive inaccuracy into three components: the distance of the population mean from the phenotypic optimum; variation of the phenotypic optimum; and variance of the population phenotype around the population mean ('population imprecision'; Armbruster *et al.*, 2004, 2009; Hansen *et al.*, 2006; Pélabon & Hansen, 2008). The population imprecision itself comprises two components: genetic imprecision (variance of the genetic values around the population mean); and developmental/environmental variance around the genotypic target (phenotypic imprecision or noise; Hansen *et al.*, 2006). Except in the brief theoretical introduction below, we do not separate empirically these last two components.

In the present study, we combine adaptive accuracy and adaptive surface approaches in an attempt to understand the causes of floral diversification, or lack of diversification, in three genera for which we have assembled data on floral morphology and pollination. We focus on a pair of traits that are easily measured and for which 'adaptive optima' (with reference to a component of fitness) can be readily hypothesized. We explore the degree to which traits track their hypothesized adaptive optima through macroevolutionary time and space. We first begin with a few comments on previous work and then give a brief introduction to the theory of adaptive accuracy, before introducing the three study systems in Materials and Methods.

Adaptive accuracy and fitness surfaces in flower pollinator 'fit'

Previous work on the adaptive accuracy of flowers has largely focused on the fit between flowers and pollinators with respect

to the location of the reward and the length of pollinator structures obtaining the reward. Darwin (1877) speculated on the evolutionary match of floral nectar spurs and the proboscides of pollinators, invoking coevolution. Indeed, some of the best evidence for the coevolution of floral spurs or tube and pollinator appendage length has come from comparisons among conspecific populations and congeneric species (Steiner & Whitehead, 1990, 1991; Johnson & Steiner, 1997). There are also a few studies of relative fitness and phenotypic selection within populations, showing higher fitness in floral phenotypes with nectar spurs or tubes that match the pollinators' proboscides (Nilsson, 1988; Maad & Alexandersson, 2004).

The adaptive surface model of natural selection has been applied to components of fitness influenced by floral morphology in two ways. One is to estimate the shape of the fitness component landscape by relating individual component fitness estimates to within-population variation in floral morphology (O'Connell & Johnston, 1998; Maad, 2000). The second approach, and that used here, is to estimate the fitness component landscape theoretically and test it against the observed distribution of populations and species in morphological space (Armbruster, 1990). Previously, Armbruster (1990) found that the blossom size of numerous populations and species of *Dalechampia* had apparently adapted to fit the size of the main pollinators. Large bees generally do not visit blossoms with small amounts of reward for energetic reasons (Heinrich & Raven, 1972; Armbruster, 1984), and small-reward blossoms must therefore have fertile structures, i.e. anthers and stigma, sufficiently close to the reward (nectar, oil, resin, etc.) in order to contact the pollinator when the latter collects the reward. The populations and species of *Dalechampia* evaluated in this study appeared largely to have evolved mean floral values close to the predicted optima (but see Hansen *et al.*, 2000).

In the present study, we consider an adaptive landscape not considered in previous studies of species divergence: the bivariate adaptive surface governing the accuracy of pollen placement on pollinators in relation to stigma contact with pollinators (a component of fitness; see Grant, 1971; Stebbins, 1974; Faegri & van der Pijl, 1979). This fitness component function can be paraphrased simply as: the expectation that anthers and stigmas contact the pollinators in the same place, which will be generally reflected as an isometric (45° slope) relationship between stamen length and style length (or equivalent measurements; see Conner & Via, 1993; Conner, 1997; Armbruster *et al.*, 2004, 2009).

It should be emphasized that adaptive surface analysis is a heuristic tool. We are not usually able to assess total lifetime fitness, but instead only components of fitness. Hence, our surface will apply to rates of pollen dispersal, pollen arrival or seed set, whilst ignoring survival, herbivory, seed predation and often offspring quality. One can derive more complete surfaces, but they are still likely to be simplified models. This is an especially important admission when applying adaptive surfaces to more than one species; different species are likely to experience

a diversity of conflicting selective pressures and constraints, and hence the surface that governs all the study species will necessarily be related to a restricted set of fitness components.

Logic of the adaptive surface of the stamen–stigma ‘fit’

A functional analysis of pollination mechanics in the context of adaptive accuracy theory leads to the expectation that the highest male function fitness accrues to individuals that place pollen on pollinators where stigmas are most likely to contact them. On the female side, highest fitness accrues to individuals whose stigmas contact pollinators where the pollen is most likely to be. In terms of floral morphology, this translates into the expectation that the ‘reward–anther distance’ [the distance between the anther and the reward or the floral constriction (‘throat’) that stops the pollinator from getting any closer to the reward] will match the mean ‘reward–stigma distance’ (the distance between the stigma and the reward or throat). This is the male component of pollination fitness. The reverse (stigma match to anther position) corresponds to the female component of pollination fitness. In terms of adaptive accuracy theory, the maximum male accuracy (for a genotype or a population; see Hansen *et al.*, 2006) is achieved when the mean reward–anther distance equals the mean reward–stigma distance (high ‘mean optimality’) and there is low variance about that mean (high floral ‘precision’). In turn, maximum female accuracy (for a genotype or population) is achieved by the mean reward–stigma distance equalling the mean reward–anther distance (high mean optimality) and having low variance about that mean (high precision; Fig. 1). We first test these simple predictions and then interpret deviations from the expected patterns in the light of constraints and conflicting selective pressures.

It should be noted that several simplifying assumptions are embedded in this conceptual model of floral fitness, and thus our model may not apply to all systems and circumstances. We assume that fitness rises monotonically with: (1) increasing amounts of pollen arriving on stigmas; and (2) increasing amounts of pollen being placed in the ‘right place’ on legitimate pollinators (and then dispersed to conspecific stigmas). We thus ignore possible negative effects of excess conspecific pollen on stigmas, but assume, instead, that fitness is enhanced by intensified pollen competition even after seed production has been maximized. We also ignore possible interactions and frequency dependence, such as selection for longer styles, when pollen competition is of intermediate intensity (see Mulcahy, 1983; Armbruster *et al.*, 1995; Armbruster, 1996; Lankinen & Skogsmyr, 2001). We also ignore complexities related to saturation of the pollen-carrying capacity of the pollinator. Under certain circumstances, it may be advantageous to place pollen on the pollinator somewhere other than the place most frequently used by other conspecific flowers, because the site is already saturated and new pollen falls off (although layering may be more common; see Harder & Wilson, 1998).

Fig. 1 The postulated axis of the adaptive ridge governing the accuracy of pollination is the isometric line ($y = x$) relating the location of pollen placement to the expected location of stigma contact with the pollinator. It also relates the location of stigma contact to the expected location of pollen placement on pollinators. Population means may lie close to, or far from, the ridge (optimality of the mean), and individuals in a population may be close to, or far from, the mean (population precision). Ellipsoids marked P1–P5 represent the spread of individual values with five populations. The broken lines indicate parallel contours of the slope.

We ignore for the moment the tendency of some flowers to place pollen in several places on pollinators (as a result of either variation among flowers in a population or among stamens within each flower), creating ‘horizontal heterogeneity’ or structure (Harder & Wilson, 1998), which may select for multiple stigma positions and/or increased variance [Armbruster *et al.*, 2009; see the extensive literature on accuracy in heterostylous flowers (for example, Sanchez *et al.*, 2008, and the studies cited therein)]. These situations may require modifications of the models presented here, but because the study systems we examined do not appear to show variation of this sort, we do not explore these issues further.

At the population and species levels, the above considerations lead to the expectation of correlated divergence of reward–anther and reward–stigma distances. This is because this type of interaction between traits generates correlational selection: selection on reward–anther distance is influenced by the value of the reward–stigma distance in a population, and vice versa. Thus, selection on each trait will be influenced by the value of the other, setting up trait covariance across populations (and species) as each achieves its adaptive combination of means. However, because selection is related to differential reproduction, but not differential survival, this correlational selection generates covariance only among, not within, populations (see Wallace, 1975; Endler, 1986, 1995; and Armbruster & Schwaegerle, 1996 for further discussions). This relationship should be reflected in populations and species falling out

on an isometric (45°) line passing near the origin. In other words, we expect populations and species to have diverged along an adaptive ridge running on a 45° diagonal across the bivariate morphological space defined by the anther–reward and stigma–reward distances (Fig. 1).

In considering adaptive covariance as a source of trait correlation, we need to be cognizant of the fact that factors other than correlational natural selection can generate covariance between traits (for example, Lande & Arnold, 1983; Armbruster, 1991; Armbruster & Schwaegerle, 1996). First, style and stamen length may be genetically correlated because of overlapping genetic–developmental control systems (pleiotropy; for example, Conner, 1997, 2002). Second, selection for larger overall flower size by pollinators could generate among-population and among-species correlations between floral traits, even if they are genetically independent (Armbruster & Schwaegerle, 1996). We consider these alternative hypotheses in the context of the data we present below.

Departures from the optimum

There have been extensive discussions as to why organisms might exhibit genetic load or maladaptation and depart from their selective optima (for example, Bradshaw, 1991; Williams, 1992; Orzack & Sober, 1994a,b; Thompson *et al.*, 2002; Hansen & Houle, 2004; see also Nesse, 2005). Contributing factors include genetic factors, such as drift, gene flow, pleiotropy and lack of genetic variation, as well as natural selective factors, such as lag in response to rapidly changing species interactions, among others. In the context of the measurement of adaptive accuracy, we wish to evaluate the roles of two genetic factors: floral integration (the tendency of floral structures to be fused and/or their variation be correlated), or lack thereof, wherein independent random variation of floral parts decreases precision and/or mean optimality; and developmental and genetic ‘constraints’ (for example, pleiotropy).

Lack of floral integration may limit a population’s ability to stay perched on the adaptive ridge, and this may drive selection for increased integration and reductions in the number of floral parts (Stebbins, 1951, 1974; Armbruster *et al.*, 2004). Other possible genetic effects include developmental relationships and genetic correlations that preclude independent evolutionary optimization of the reward–anther and reward–stigma distances (Armbruster & Schwaegerle, 1996; Schluter, 1996; Hansen *et al.*, 2003a). Comparison of the population mean with species mean conformance with the postulated adaptive surface may reveal the effect of genetic/developmental constraints. This is because genetic constraints will usually have stronger effects on covariation within species than among species (Endler, 1986, 1995; Armbruster, 1991; Armbruster & Schwaegerle, 1996), because the **G** matrix is itself a potentially evolving ‘trait’ at the level of populations and species (Lande, 1980; Turelli, 1988; Jones *et al.*, 2003, 2004; Revell, 2007; Polly, 2008; Arnold *et al.*, 2008).

With respect to other components of fitness, we wish to assess the role of possible conflicting selective pressures in driving departure from the modelled adaptive optima. Most traits are influenced by several selective pressures and, when these involve trade-offs, it is usually impossible to respond optimally to all (Schluter *et al.*, 1991; Strauss & Irwin, 2004), leading to adaptive compromise (see Armbruster, 1996, 2001 for floral examples). One little studied, but striking, floral example is the conflict between selection for increased outcrossing in self-compatible species that are not dichogamous (sexual functions not temporally separated) and selection for placing pollen in the same place on pollinators contacted by stigmas. The former favours herkogamy (spatial separation of anthers and stigmas), but this may often reduce the correspondence in the points of anther and stigma contact with pollinators, hence reducing the mean optimality in our analysis.

Interestingly, there are at least three possible routes of escape from the trade-off between accuracy and herkogamy (outcrossing) in monomorphic flowers (the situation differs for heterostylous flowers; see Discussion), and we wished to explore their effects on floral accuracy and its components. One ‘escape route’ is being a self-pollinator that ‘tolerates’ inbreeding (although this is more likely to preclude the conflict rather than be an escape from it). A second route is to escape in time by segregating male and female functions temporally (dichogamy). The third route is to achieve herkogamy whilst maintaining accurate fit with pollinators by escaping into higher dimensional space. This works in some flowers by having reward–anther separation in one dimension and reward–stigma separation in another. In this situation, the separation between anthers and stigmas is greater than the difference between the two distances from the reward (see below for a further explanation).

Theoretical basis of adaptive accuracy

We present here only a brief précis of the theory of adaptive accuracy. More detailed accounts are given in the treatments by Hansen *et al.* (2006) and Armbruster *et al.* (2009). Although maladaptation and inaccuracy are logically measured on individuals, they are also properties of populations, and it is the latter application that is used here to study population and species divergence.

Consider first the dispersal of pollen to other stigmas, as determined by the deposition of pollen on pollinators (i.e. the male function). Let θ be a random variable with a specified distribution, representing the optimal position of stigmas in the population relative to the landmark. [The landmark is usually the reward or the perianth restriction (‘throat’) that stops the pollinator getting any closer to the reward.] We can think of deviation from the optimum as decreasing fitness and therefore being subject to selection. Selection operating in the context of a single population thus has components relating to the various causes of phenotypic deviation from the optimum, and maladaptation at the population level is the sum of these components:

$$s(E[z] - E[\theta])^2 + s\text{Var}[Z_t] + s\text{Var}[\theta] + sE[V_d] \quad \text{Eqn 1}$$

[s , strength of stabilizing selection; E , expected value of the variable in the following brackets; z , the observed phenotype; θ , optimal phenotype; $\text{Var}[Z_t]$, variance in the genotypic target (the target is the expected phenotype produced by a genotype); $\text{Var}[\theta]$, variance in the optimum; V_d , variance in the phenotype around the genotypic target as a result of environmental variation and developmental noise].

There are thus four components of inaccuracy to consider when assessing population properties. These four components can be operationalized as: (1) the 'bias', $E[z] - E_\theta[\theta]$, which is measured as the difference between the population trait mean and the population optimum; (2) the variance of the fitness optimum, $\text{Var}[\theta]$, which is measured as the population variance in the optimum; (3) the variance in the genotypic target, $\text{Var}[Z_t]$, which is measured as the population variance of genotype means; and (4) the phenotypic imprecision resulting from developmental noise and environmental variance V_d , which is measured as the within-plant variance for the focal trait (for example, across flowers on a plant), and $E[V_d]$ is treated as the mean within-plant variance for the population. However, for population studies in the field, it is useful to pool terms (3) and (4) and estimate them jointly as the within-population phenotypic variance of the trait. This leads to a simplified measure of inaccuracy:

$$\text{Inaccuracy} = (E[z] - \theta)^2 + \text{Var}[\theta] + \text{Var}[z] \quad \text{Eqn 2}$$

In other words:

$$\text{Inaccuracy} = (\text{Population Trait Mean} - \text{Optimum})^2 + \text{Variance of Optimum} + \text{Population Imprecision}$$

This measure of inaccuracy and its components have units equal to the trait units squared. For comparison across species and traits, they can be standardized by dividing by the trait mean squared. When this is performed, their numerical value can be interpreted as the percentage reduction in fitness, when the mean standardized selection coefficient s in Eqn 1 is equal to unity (i.e. $sE[z]^2 = 1$).

Materials and Methods

Study systems

We compared patterns of floral optimality and accuracy within and among three genera for which we have extensive morphological datasets and phylogenetic information (two of three taxa). These three systems are drawn from distantly related families, and hence represent a broad sample of angiosperms. They also represent a broad range of types of floral organization. *Dalechampia* (Rosidae: Euphorbiaceae) has pseudanthial, functionally bisexual blossoms as pollination units; these comprise

unisexual flowers and hence have low structural integration (blossom parts are developmentally more independent and/or show less fusion than parts of a single flower) compared with the other two genera (Webster & Webster, 1972; Armbruster, 1988, 1993; Armbruster *et al.*, 2004). *Collinsia* and *Tonella* (Asteridae: Plantaginaceae: Collinsieae) have flowers as pollination units, and these have an intermediate level of structural integration by fusion within (connation: synsepaly, sympetaly, syncarpely) and among (adnation: epipetalous stamens; Armbruster *et al.*, 2002, 2004) whorls. *Styloidium* (Asteridae: Stylidiaceae) has flowers as pollination units, and these have an even greater level of structural integration by within-whorl fusion (synsepaly, sympetaly, syncarpely) and among-whorl fusion (complete adnation of staminate and pistillate tissues; Armbruster *et al.*, 1994, 2004).

Dalechampia is a clade of *c.* 120 species of mostly perennial vines, distributed throughout most of the lowland tropics. The bilaterally symmetrical, laterally oriented, blossom inflorescence (pseudanthium) usually comprises 10–15 staminate flowers, three pistillate flowers and a gland that, in most species, secretes resin (*c.* 100 species) or fragrance (three species). These parts are subtended by two, usually showy, bracts. The reward and all floral parts are fully exposed when the bracts are open, but pollinators generally orient themselves consistently on the bilaterally arranged flowers whilst collecting resin or fragrance. Pollination of most species is by resin-collecting, female bees, which use resin in nest construction, or fragrance-collecting, male bees, which probably use fragrances to attract females (Armbruster, 1993). We measured several floral size and orientation traits on flowers from usually 5–45 plants per population, 1–20 populations per species and 35 species with digital or dial callipers precise to 0.01 mm.

Collinsia and its close relative, *Tonella*, form a clade (tribe Collinsieae) of *c.* 25 annual species, primarily of temperate western North America (Armbruster *et al.*, 2002). The flowers are zygomorphic (bilateral), with a landing platform formed by the lower lip and a banner formed by the upper lip. The four stamens and style are enclosed in a keel-like fold of the lower lip, and exposed only when a nectar-seeking bee of sufficient size lands on the flower. Pollination is by long-tongued, nectar-feeding bees (which may also collect pollen; Armbruster *et al.*, 2002). We measured flowers on 5–20 plants per population, one to eight populations per species and 24 species with digital or dial callipers precise to 0.01 mm.

Styloidium contains over 250 species of herb, perennial rosette plants and small shrubs, most of which are endemic to Australia. The flowers are zygomorphic and characterized by the fusion of staminate and pistillate tissues into a motile, protandrous column. Pollination is by nectar-feeding bee flies (Bombyliidae) and small solitary bees which, on contacting the trigger-point whilst foraging for nectar, cause the column to spring forward to place pollen on, or pick it from, the back, side or venter of the pollinating insect (Armbruster *et al.*, 1994). The column 'reloads' to the original position in *c.* 30 min, and

has the ability to repeat this action numerous times (40+ times) in the c. 3–5 d life of a flower. The flowers place pollen on pollinators in the first 1–2 d of receptivity and then pick up pollen in the same way in the final 1–2 d. We measured flowers on 5–10 plants per population, 1–12 populations per species and 31 species with digital or dial callipers precise to 0.01 mm.

Measurements and analysis

In this analysis of floral accuracy and pollination fitness, we restrict our attention to the match of stigma position to the site of expected pollen deposition on pollinators, and the match of anther position to the expected site of stigma contact with pollinators. We used three study systems in which pollinators are largely immobile after landing on the flower. This allows us to use floral measurements (the distance between the floral landmark, e.g. resin gland or throat of floral tube, and the anthers or stigmas) to predict quite closely the sites of pollen placement and pick up, respectively, on the pollinator (see Armbruster *et al.*, 2009). This is not the case if pollinators crawl around on the flowers.

Following Eqn 2, we calculated the fitness decrement resulting from 'floral inaccuracy' as:

$$(\text{Mean Reward–Anther Distance} - \text{Mean Reward–Stigma Distance})^2 + \text{Variance in Reward–Stigma Distance ('variance of optimum')} + \text{Variance in Reward–Anther Distance ('population precision')} \quad \text{Eqn 3}$$

This is the absolute inaccuracy of both male and female functions in this particular system, because the variance of the optimum for male inaccuracy is the variance of the female trait, and vice versa (this is not a general property, however). In other words, although Eqn 3 is actually male inaccuracy, it is obviously equivalent to the equation for female inaccuracy:

$$(\text{Mean Reward–Stigma Distance} - \text{Mean Reward–Anther Distance})^2 + \text{Variance in Reward–Anther Distance (variance of optimum)} + \text{Variance in Reward–Stigma Distance (population precision)} \quad \text{Eqn 4}$$

As a result of this equivalence for the traits under study here, we treat the value as the 'joint floral inaccuracy' of both male and female functions.

Because the variance of morphological measurements usually scales with the trait means, we scaled inaccuracy calculations before making comparisons between species and between study systems with flowers of different sizes and shapes. We scaled the joint inaccuracy with the product of the trait means, which is, in fact, the square of the geometric mean of the two traits (see Sokal & Rohlf, 1981; Hansen *et al.*, 2003b). Such scaling is desirable because it conserves the additive properties of the variance components, a property that coefficients of variation (CVs) do not have.

Because, in this system, the variance of the optimum is the same as the variance of the alternative target trait (hence the equivalence above), we also wished to assess the independent contributions of male (staminate) and female (pistillate) functions to adaptive inaccuracy. We therefore calculated 'pure' male inaccuracy at the population level as:

$$(\text{Mean Reward–Anther Distance} - \text{Mean Reward–Stigma Distance})^2 + \text{Variance in Reward–Anther Distance} \quad \text{Eqn 5}$$

The 'pure' female inaccuracy at the population level was thus defined as:

$$(\text{Mean Reward–Stigma Distance} - \text{Mean Reward–Anther Distance})^2 + \text{Variance in Reward–Stigma Distance} \quad \text{Eqn 6}$$

We calculated the above components of inaccuracy and, for purposes of comparison, scaled them to the square of the mean trait values and converted them to percentages. This allows all components to be compared among study systems and traits, whilst maintaining their additive properties. When scaled in this way, the imprecision component reduces to I , the mean-squared scaled phenotypic population variance. $I (=CV^2)$ has theoretical advantages related to additivity and interpretation as trait evolvability (see Hansen *et al.*, 2003b; Hansen & Houle, 2008). Inaccuracies and mean departure from optimality were also scaled to the square of the trait mean and converted to percentages for comparisons across traits, populations and study systems.

Our analysis of interpopulation and interspecific data took several approaches. First, we wished to test the idea that there is a fitness surface governing the interaction of reward–stigma and reward–anther distances across multiple species. We examined the relative positions of anthers and stigmas, treating them as bivariate morphological space. We hypothesized that maximum fitness is a positive, isometric adaptive ridge passing through (or near) the origin. We tested this proposition by mapping population and species means onto the hypothesized adaptive surface for three distantly related genera. We then considered the adaptive accuracy of a sample of species drawn from these genera, assessing the relative contributions of imprecision and mean departure from the optimum to inaccuracy. We attempted to discover reasons for local departures from the adaptive ridge, considering floral integration and precision, genetic constraints and conflicting selective pressures. We also tried to refine our understanding of the shape of the adaptive surface, specifically whether the ridge was broad or narrow.

We analysed patterns of species divergence by relating population means and variances to the hypothesized adaptive surface, an adaptive ridge depicting the trajectory of highest accuracy. As noted above, this surface is hypothesized to govern the pollination component of fitness, but not necessarily total fitness, except under a *ceteris paribus* assumption ('other things equal' is a simplifying assumption in these analyses). The

adaptive hypothesis is based on empirical observations made on these three study systems (Armbruster, 1988, 1990; Armbruster *et al.*, 1994, 2002) and the simple logic that, for pollen to reach a stigma, it must be placed in a location on the pollinator that touches stigmas on subsequent visits to other flowers. Similarly, for a stigma to receive pollen, it must contact the pollinator in the location in which pollen has been previously placed by other flowers (see Armbruster *et al.*, 2004, 2009).

Numerical characterization and comparisons were based on correlation statistics, multiple regression, path analysis and calculation of several less well-known evolutionary parameters, such as evolvability and conditional correlation (see Hansen *et al.*, 2003a,b). Most population/species comparisons were calculated from the sum of all measurements of that trait for all populations (i.e. weighted means rather than means of genotype means). Species means, however, were calculated from the population means without weighting. Analyses of population means across species implicitly ignored phylogenetic structure and possible heterogeneity in slopes of relationships among species and at the population and species levels (see Armbruster, 1988, 1991; Bell, 1989). However, we felt that this problem was minor because, with a few exceptions, only a few populations were sampled per species.

In order to test whether the correlation between the gland–stigma distance (GSD) and gland–anther distance (GAD) is caused by a spurious relationship with gland area (GA) (and selection by bees for small or large GA), we calculated conditional correlations following the method of Hansen *et al.* (2003a). We first used maximum likelihood estimators (dividing by n not $n - 1$) to compute the variance–covariance matrix for *Dalechampia* population means with complete data for GA, GSD and GAD. We then computed the variance matrix of GAD and GSD conditional on GA using the following relationship (see Hansen *et al.*, 2003a):

$$V_{y|x} = V_y - V_{yx}V_x^{-1}V_{xy} \quad \text{Eqn 7}$$

(V_x^{-1} , inverse of V_x ; V_{yx} , covariance matrix between y and x). In this analysis, $y = \{\text{GAD, GSD}\}$, $x = \text{GA}$ and $V_{yx} = \{\text{Cov}[\text{GAD, GA}], \text{Cov}[\text{GSD, GA}]\}$.

Hypothesis testing of statistical analyses of interspecific trends in *Dalechampia* and *Collinsia* was based on phylogenetically informed independent contrasts implemented in ‘Comparative Analysis by Independent Contrasts’ (CAIC) (Felsenstein, 1985; Purvis & Rambaut, 1995) using published or in-press molecular phylogenies (see Armbruster & Baldwin, 1998; Armbruster *et al.*, 2002). It was not possible to assess the phylogenetic contribution to the trait correlations in *Stylidium* because of the absence of an independent phylogenetic estimate, although this was probably not a serious problem because of the apparent extreme evolutionary lability of column length and the tight relationship between male and female functions (see Armbruster *et al.*, 1994).

Results

Tests of the hypothesized adaptive surface governing the stamen–stigma ‘fit’

As expected, the population and species means of all three study groups fell near the crest of the hypothesized adaptive ridge (Fig. 2). The tightness of the fit is indicated by the R^2 values, which ranged from 0.615 in *Collinsia* and 0.723 in *Dalechampia* to near 1.0 in *Stylidium*.

The fit of populations and species to an isometric line is only a weak test of the adaptive ridge hypothesis, in so far as there are other possible reasons for such a relationship. One possible alternative is that larger pollinators select for larger flowers (and floral structures) than do smaller pollinators, and that this relationship has generated spurious covariance (in the path analytical sense; Li, 1975) between reward–anther and reward–stigma distances. We were able to test this idea in *Dalechampia* by assessing the role of GA (a determinant of pollinator size and hence a reasonable proxy for it; Armbruster, 1988) vs GSD as potential ‘determinants’ of GAD. If all the covariance between GSD and GAD were explained by the effect of phenotypic correlations with GA, selection on gland size by pollinators, rather than selection for accuracy, would explain the observed GAD–GSD covariance across populations and species. We tested this by computing the covariance of GAD and GSD conditionally on GA (Hansen *et al.*, 2003a) across population means. Although the conditioning reduced the covariance from 5.40 to 1.18 mm², underscoring the importance of gland size, the trait variances were also reduced, and a strong correlation of 0.71 remained after conditioning on GA. This is lower than the unconditional correlation of 0.89, but still shows that covariance between anthers and stigma is caused by more than overall blossom size. A path diagram illustrates this relationship using partial regression statistics, showing that even the partial effect of GAD on GSD is quite strong (Fig. 3). (It should be noted that whether GAD or GSD is used as the dependent variable in this exercise is purely arbitrary; swapping around the dependent variables reduced even further the importance of GA.)

We also evaluated the strengths and trajectories of the hypothesized adaptive covariance relationship in comparison with other intertrait relationships in order to assess further the likelihood that the expected among-species relationship is simply the result of general genetic or phenotypic covariation in the size of floral traits, rather than a fit to an adaptive landscape. For *Dalechampia* species means (mean scaled to reflect proportional slopes), the hypothesized adaptive correlation between GSD and GAD was larger ($r = 0.86$) than all but one of the five other trait correlations (mean $r = 0.72 \pm 0.058$), as expected. Similarly, the intercept of the GSD–GAD trajectory was much closer to zero than for any other trait combination (0.0003 vs mean intercept value of 0.47 ± 0.043), as expected. The slope was closer to 1.0 than any other trait combination (0.9997 vs 0.53 ± 0.043).

Fig. 2 Bivariate plots of population means relative to the hypothesized adaptive ridge for the three study systems. Broken lines indicate the isometric line hypothesized to be the adaptive ridge governing the fitness response to both intraspecific and interspecific variation in anther–reward and stigma–reward distances. Each point represents a population mean for the two traits. (a) Seventy-four population means from 28 species of *Dalechampia* (Euphorbiaceae) in relation to the hypothesized adaptive ridge governing the distance between the resin gland and the anthers (GAD) and the stigmas (GSD); square points and regression represent populations of one species, *D. scandens*. (b) Thirty-one population means from 15 species of the monophyletic tribe Collinsieae (Plantaginaceae) in relation to the hypothesized adaptive ridge governing the distance between the floral throat and the anthers and the stigmas. (c) Twenty-one population means from 11 species of *Stylidium* (Stylidiaceae) in relation to the hypothesized adaptive ridge governing the length of the column in the staminate and pistillate phases.

Fig. 3 A path diagram illustrating the large partial effect of the gland–stigma distance (GSD) on gland–anther distance (GAD) after controlling for the effects of gland area. The choice of GAD as the dependent variable was arbitrary, but the same pattern is seen if GSD is used as the dependent variable. The data analysed were the population means; the numbers are the standardized path coefficients (= standardized partial regression coefficients), which vary from -1 to $+1$, with values near zero indicating no effect.

For *Collinsia* species means, the hypothesized adaptive correlation between mean-scaled style and stamen lengths ($r = 0.89$) was larger than the mean of the nine other trait correlations (mean $r = 0.81 \pm 0.03$), as expected. The expectation was that the intercept (a) would be closer to zero than for the other traits, and this was indeed the case: $a = 0.10$ for stamen–style length vs a mean of $0.24 (\pm 0.07)$ for other trait combinations. The expectation for the slope of the stamen–style length relationship was that it would be closer to 1.0 than were other trait combinations; this was indeed the case: 0.90 vs $0.76 (\pm 0.07)$. For the mean-scaled *Stylidium* species means, the adaptive correlation was estimated as 1.0, which was much larger than any other trait correlation (mean $r = 0.42 \pm 0.084$). The observed intercept was approximately zero, as expected, vs a mean for other trait combinations of 0.63 ± 0.061 . The regression slope was close to 1.0, as expected, vs an average of 0.37 ± 0.061 for the other trait combinations. Together, these data indicate that in none of the three study systems can the fit of anther and stigma distances to the hypothesized ridge be explained as a pure allometric response to variation in overall flower size.

Adaptive accuracy and precision

The mean $\pm$ standard error (SE), joint mean-product-scaled, blossom inaccuracy (where joint floral inaccuracies are scaled to the product of the mean GSD and mean GAD) of 74 populations of 28 species of *Dalechampia* was $14.27 \pm 1.30\%$. The corresponding mean, mean-product-scaled, floral inaccuracy of 31 populations of 15 species of the Collinsieae clade (*Collinsia* and *Tonella*) was $16.56 \pm 1.17\%$. These values would cause a substantial reduction in fitness if stabilizing selection was strong ($sE[z]^2 \sim 1$ would cause a similar percentage reduction in fitness). By contrast, the mean, mean-product-scaled, floral inaccuracy of 21 populations of 11 species of *Stylidium* was only $0.98 \pm 0.21\%$ (Table 1).

The three contributors to joint floral inaccuracy differed in their importance across the three systems. Mean departure from the optimum was by far the most important factor in

Fig. 4 Path diagrams illustrating the relative, unique contributions of stamen (male) imprecision, stigma (female) imprecision and population mean departure from the optimum to the joint floral inaccuracy in: (a) 74 populations in 28 species of *Dalechampia* (Euphorbiaceae); (b) 31 populations in 15 species of the tribe Collinsieae (Plantaginaceae); and (c) 21 populations in 11 species of *Stylidium* (Stylidiaceae). Joint floral inaccuracy was measured as the sum of the variances in stigma position and stamen position plus the square of the difference between stigma and stamen position. Numbers are standardized path coefficients (= standardized partial regression coefficients), which vary from -1 to +1, with values near zero meaning no effect. All variables were mean-squared scaled and hence are unit-less percentages.

Fig. 5 The relationship between mean, mean-squared-scaled male inaccuracy and mean female imprecision ($I = CV^2$) in 31 populations in 15 species of the tribe Collinsieae (Plantaginaceae).

Dalechampia, although it was moderately correlated with male imprecision (Fig. 4a). In Collinsieae, mean departure from the optimum and female imprecision were roughly equally important. Mean departure from the optimum and female imprecision were strongly correlated (Fig. 4b). The joint floral accuracy of *Stylidium* spp. was the result of imprecision only (Fig. 4c), because mean deviation from the optimum was estimated as zero in all species.

Male and female accuracy and precision compared The joint floral inaccuracies of *Dalechampia* and Collinsieae were quite similar. However, when male and female components were separated, we detected a more important contribution of female imprecision in *Collinsia* (i.e. variance in stigma position) compared with male imprecision (Fig. 4b). By contrast, these two factors were about equally important in *Dalechampia* (Fig. 4a).

Male and female inaccuracies were generally correlated because they both contain the same term: mean optimality. However, male and female precisions are independent and can be compared meaningfully. Surprisingly, they were not strongly correlated in either *Dalechampia* ($r = 0.12$) or *Collinsia* ($r = -0.13$). *Stylidium* could not be assessed because male and female precisions were not measured independently.

Another expectation was that male inaccuracy (not including the optimum variance term) would track female precision (optimum variance) because, if stigma positions are highly variable, there would be relaxed selection for male accuracy. As discussed below, stigma positions are sometimes subject to conflicting selective forces and constraints, which may lead to significant imprecision. Male inaccuracy did not track female imprecision in *Dalechampia* ($r = 0.005$), but did so closely in *Collinsia* ($r = 0.826$; independent contrast $P < 0.01$; Fig. 5). A similar trend is suggested by comparing the study systems:

Stylidium species had lower mean female imprecision ($6.95 \pm 2.25\%$) and lower mean male inaccuracies ($0.98 \pm 0.58\%$), whereas *Dalechampia* and *Collinsia* had higher female imprecision ($17.3 \pm 6.3\%$ and $29.9 \pm 13.7\%$, respectively) and higher male inaccuracies ($13.7 \pm 19.6\%$ and $6.2 \pm 5.3\%$, respectively).

Although we would similarly expect female inaccuracy to be influenced by male imprecision, there was no detectable relationship in either *Dalechampia* or *Collinsia*. We might also expect male precision to track female accuracy (inaccurately positioned stigmas might select for lower precision in stamen position), although no such trend was detectable. Similarly, female precision might track male accuracy, and this relationship was detected in *Collinsia*, but it cannot be distinguished from the relationship shown in Fig. 5 (with axes inverted). No such relationship was detected in *Dalechampia*.

Causes of mean departure from the optimum

Genetic and mean developmental constraints Random variations in floral parts that lack integration may increase both imprecision and departure from the optimum. This may explain the low accuracy of *Dalechampia* spp. and *Collinsia* spp. relative to *Stylidium* spp. *Dalechampia* have male flowers functioning as stamens and female flowers functioning as pistils in a pseudanthial inflorescence (blossom), and hence the fertile structures are not only unfused, they are in different flowers, and only secondarily coordinated (that is at the level of the inflorescence rather than the flower). In Collinsieae, both sexual functions are in single perfect flowers, which have fused (connate) petals and staminal filaments adnate (fused) to the base of the corolla (epipetalous stamens). Both groups have comparatively high inaccuracies, although it is surprising that *Dalechampia* is as accurate as *Collinsia*, given its low level of structural integration. *Collinsia* suffers from another mechanical genetic 'constraint': its stamens are enclosed in a narrow keel (in so far as the keel is adaptive, this is ultimately a selective trade-off). Thus, *Collinsia* spp. cannot escape from the herkogamy–accuracy trade-off by using higher dimensional space (see below). By contrast, *Stylidium* species have male and female tissues fused into a single structure. The flowers are thus highly coordinated in contacting the pollinators in a consistent place with both the anthers and stigmas sequentially. The remarkably high accuracy and precision of *Stylidium* flowers is at least partly a result of this integration.

Comparison of the population mean vs species mean conformance to the postulated adaptive surface may reveal the effect of genetic/developmental constraints. This is because genetic constraints will usually have stronger effects on covariation within species than among species, because, although the G matrix is itself a potentially evolving trait, divergence may require considerable time. The trajectory of among-species covariation of GAD and GSD of *Dalechampia* appears to fit to the isometric adaptive ridge, as expected (Fig. 2a). However, a sample of South American populations of *D. scandens*

Fig. 6 The relationship between gland–anther distance (GAD) and gland–stigma distance (GSD) across 18 Mexican populations of two hypothesized cryptic species in the *Dalechampia scandens* complex, where the two clusters of points in morphometric space represent the two hypothesized cryptic species. (The data presented here were not included in Fig. 2 or related analyses.) It should be noted that one point (in the broken circle) did not cluster with any other points and was excluded from later analyses. The open circles are the means of the two cryptic species (left and right clusters). Populations within a cryptic species do not appear to track the hypothesized adaptive ridge (straight broken line), although the means of the two cryptic species may do so. The parameter estimates for this relationship are based on regression of all 18 population means.

appears to follow a trajectory ($b = 0.54$) closer to the genetic regression ($b = 0.67$; measured in one population) than the hypothesized adaptive trajectory ($b = 1.0$) (Fig. 2a; Hansen *et al.*, 2003b). The analysis of two cryptic species of the *D. scandens* complex in Mexico shows a poor fit of populations within each cluster (hypothesized cryptic species) to the adaptive ridge. The predicted regression values were: intercept = 0, slope = 1.0, $R^2 \approx 1.0$, but the observed values were 1.77, 0.36, 0.31 (left cluster, facultative–selfing species), and 5.99, –0.09, 0.004 (right cluster, facultative–outcrossing species), respectively. By contrast, the means of the two subspecies conformed to the adaptive ridge reasonably well (Fig. 6).

Conflicting selective pressures Selection for increased outcrossing will favour herkogamy in self-compatible species that are not dichogamous (sexual functions separated temporally), and response to this selection may reduce the optimality of the mean. This relationship can be examined by comparing the optimality scores of self-compatible species that are facultative selfers vs facultative outcrossers, because facultative selfers are presumably not under strong selection for herkogamy (or may even experience selection against herkogamy), whereas facultative and obligate outcrossers presumably are.

The predicted relationship in *Dalechampia* is for species with high selfing rates and little or no herkogamy to have higher optimality scores and potentially higher accuracies. We used the anther–stigma distance (ASD) as a proxy for the

outcrossing rate (see Armbruster, 1988), but did not find any relationship between this measure and optimality. In *Collinsia*, however, there was a significant positive relationship between the

Fig. 7 Response of the mean departure from the optimum (circles) and imprecision stamen (triangles) and style (squares) length (l) to variation in the outcrossing index (calculated as the sum of the relativized corolla length and the relativized time of self-pollination, where near four is most highly outcrossing and near zero is most highly selfing; see Armbruster *et al.*, 2002) across 31 populations of 22 species of *Collinsieae* (Plantaginaceae). See text for statistical analyses.

estimated outcrossing rate and mean-squared scaled departure from the optimum ($r = 0.75$; independent contrast $P < 0.001$) and female imprecision ($r = 0.63$; independent contrast $P < 0.01$), but no clear relationship with male imprecision ($r = -0.30$, independent contrast $P > 0.20$; Fig. 7).

Escape from the conflict of herkogamy There seems to be evidence for escape from the herkogamy–accuracy trade-off by some populations of *D. scandens*. This mechanism is best understood by examining the blossom geometry in lateral view (Fig. 8). One group of populations clusters in morphological space (left cluster; Figs 6, 9) and conforms to the blossom form depicted in Fig. 8a. By contrast, the other populations cluster to the right in Figs 6 and 9 and conform to the blossom morphology depicted in Fig. 8b. The left cluster of populations has the three stigmas and 10 staminate flowers arranged more or less in a single plane (in a lateral view, this plane is portrayed as a line; Fig. 8a). By contrast, the right cluster utilizes higher dimensional space, with the styles diverging out of the plane formed by the resin gland and staminate flowers (in lateral view, these planes appear as two diverging lines; Fig. 8b).

The optimality consequences of this geometrical difference are shown in Fig. 9. There seems to be an initial trend towards decreasing mean optimality (increasing difference between GSD and GAD) with increasing herkogamy (ASD), that is, a trade-off, in those populations with one-plane geometry (Fig. 9; left population cluster). This trade-off disappears completely in the populations with two-plane geometry (Fig. 9).

Fig. 8 Photographs and diagrammatic representations of the two different arrangements of flowers in two hypothesized cryptic species of *Dalechampia scandens*, which do not (a) and do (b) escape from the herkogamy–optimality trade-off by ‘escape’ into higher dimensional space. Symbols: A, anthers; G, resin gland; S, stigma. (a) ‘Left cluster’ populations have small anther–stigma distances and anthers and stigmas oriented in more or less the same plane (in the lateral view, a line passing through the resin gland), hence experiencing the trade-off that $ASD = GSD - GAD$, where ASD is the anther–stigma distance, GSD is the gland–stigma distance and GAD is the gland–anther distance. (b) ‘Right cluster’ populations have large anther–stigma distances and anthers and stigmas oriented in different planes (in the lateral view, lines passing through the resin gland), hence not experiencing the trade-off that $ASD = GSD - GAD$.

Fig. 9 Relationship between mean departure from optimality (as measured by the difference between the gland–stigma distance and anther–stigma distance, GSD–GAD; y-axis) and herkogamy (as measured by the anther–stigma distance, ASD; x-axis) across 17 populations of *Dalechampia scandens* in Mexico. Points are population means. There is a possible trend towards increasing departure from the optimum with increasing herkogamy in the left population cluster (broken line). However, in the right cluster, which is of type 2 geometry, departure from optimality does not increase with herkogamy, indicating escape from the optimality–herkogamy trade-off in higher dimensional space. The overall relationship of the 17 population means is indicated by the full regression line, with the parameter estimates indicated above the graph. The distinctive blossom geometries of the two population clusters are indicated by the diagrams below.

There appears to be a similar trend across species in the rest of the genus. Species with stamens and style falling out on a single plane were categorized as having ‘type 1’ geometry, and those with the styles diverging in a different dimension were categorized as having ‘type 2’. We predicted that type 2 species would have higher mean optimality for a given level of herkogamy. As predicted, type 1 species had greater mean departure from the optimum (24.66%) than type 2 (18.67%; Table 2). This higher optimality of the mean should, in turn, select for lower imprecision in both stamens (4.50 vs 5.71%) and styles (2.73 vs 4.01%), as observed (Table 1). Although the above ANOVAs were not phylogenetically informed, the scattered distribution of type 2 geometry throughout the genus indicates five to six origins of type 2 from type 1 and 0–1 reversals (Fig. 10), and hence phylogenetic pseudoreplication is probably not a serious statistical problem.

Discussion

Logic and tests of the adaptive surface of the stamen–stigma ‘fit’

The population and species means in all three study groups fell reasonably near the top of the hypothesized adaptive ridge

Table 1 Mean values (+ standard errors) of accuracy and its components for the three study systems

Taxon	Mean mean ² -scaled male inaccuracy (%)	Mean mean ² -scaled female inaccuracy (%)	Mean mean ² -scaled joint inaccuracy (%)	Mean mean ² -scaled male imprecision (I) (%)	Mean mean ² -scaled female imprecision (I) (%)	Mean mean ² -scaled departure from optimum (male and female) (%)
Formula	$(\bar{x} - \bar{y})^2 / x^2 + \text{var}_x / x^2$	$(\bar{x} - \bar{y})^2 + \text{var}_x + \text{var}_y / y^2$	$(\bar{x} - \bar{y})^2 + \text{var}_x + \text{var}_y / xy$	var_x / x^2	var_y / y^2	$(\bar{x} - \bar{y})^2 / x^2$ and $(\bar{x} - \bar{y})^2 / y^2$
<i>Dalechampia</i> (n = 74 populations)	13.67 (± 2.28)	8.70 (± 0.89)	14.27 (± 1.30)	4.21 (± 0.473)	3.40 (± 0.282)	9.96 (± 2.05) and 5.537 (± 0.888)
Collinsiae	6.15 (± 0.96)	19.29 (± 4.30)	16.56 (± 2.61)	2.30 (± 0.369)	10.75 (± 1.86)	3.90 (± 0.556) and 6.64 (± 1.27)
<i>Stylidium</i> (n = 20 populations)	0.98 (± 0.13)	0.98 (± 0.13)	0.98 (± 0.13)	0.503 (± 0.073)	0.503 (± 0.073)	0.0 and 0.0

All measures are given as percentages, scaled to the mean squared. Male inaccuracy lacks the optimum variance and is scaled to the mean stigma distance squared. Joint inaccuracy includes both anther and stigma variances and is scaled to the product of the mean anther and stigma distances.

Fig. 10 A representative pruned maximally parsimonious tree of *Dalechampia* species showing multiple shifts in blossom geometry, representing escape from the herkogamy–accuracy trade-off by exploitation of higher dimensional space. For explanation of the diagrams, see Fig. 8. The phylogenetic estimate is based on maximum parsimony analysis of combined nuclear ribosomal (ITS-1, 5.8S, ITS-2) and chloroplast (*trnK* intron) DNA sequences (Armbruster & Baldwin, 1998; B. G. Baldwin & W. S. Armbruster, unpublished).

Table 2 Classification of *Dalechampia* species into two types of blossom geometry

Blossom geometry	Type 1	Type 2
Number of species	26	17
Average mean ² -scaled deviation from optimum (%) (standard error)	24.66* (3.66)	18.67* (4.21)
Average mean ² -scaled male imprecision (%) (standard error)	5.71** (1.07)	4.50** (1.03)
Average mean ² -scaled female imprecision (%) (standard error)	4.01*** (0.57)	2.73*** (0.50)

Not all species fit neatly into these categories. Types 1 and 2 are illustrated in Fig. 8. Differences significant under the assumption of species independence (but see Fig. 10): *, $F = 64.8.0$, $P < 0.001$; **, $F = 217.8$, $P < 0.001$; ***, $F = 275.8$, $P < 0.001$.

governing the placement of pollen on and receipt from pollinators (Fig. 2). This observation supports our hypothesis that fitness is highest when reward–stigma and reward–anther distances are nearly the same, but it does not allow statistical evaluation, as $n = 3$. *Stylidium* spp. (Stylidiaceae), which have the greatest structural integration and lowest population inaccuracy values, fit much more tightly to the hypothesized ridge than *Dalechampia* spp. (Euphorbiaceae) and Collinsieae spp. (Plantaginaceae), with lower structural integration. Indeed,

the structure of *Stylidium* flowers, with fused stamen and pistil tissues, almost guarantees a good fit to the ridge, as the lengths of the two tissues are mechanically linked.

One question that arises when modelling a multispecies fitness surface as a ridge is: what determines where individual populations and species lie along the ridge? One possibility is that genetic drift and/or random speciation generates these differences (although, of course, the combinations remain adaptive). Another possibility is that the ridge is ‘bumpy’ or a cordillera of peaks (depending on how low the ‘passes’ are). Some ecological information suggests that this is probably the case for most flowers and certainly the three studied here. In pollination systems, the adaptive ridge is likely to be extremely bumpy because pollinator size (or behaviour) often has a discontinuous distribution. For example, most *Dalechampia* species are pollinated by bees of $c.$ 5.5–7.5 mm, 9–12 mm or 20–26 mm in length (Armbruster, 1988; Hansen *et al.*, 2000). This discontinuity would create a series of high and low points along the adaptive ridge: high where both GAD and GSD match and occupied pollinator size class, and low where they do not. We would expect local high points along this ridge to fall out roughly at GAD = GSD = 3–5 mm (touching the abdomen of *Trigona* or *Hypanthidium*), GAD = GSD = 5–7 mm (touching the thorax or abdomen of *Euglossa* spp.), GAD = GSD = 8–14 mm (touching the thorax or abdomen of *Eufriesea* spp.)

and GAD = GSD = 16–22 mm (touching the thorax or abdomen of *Eulaema* spp.). Indeed, the distribution of GAD and GSD across species shows peaks and troughs in their frequency distributions (see GSD, fig 2 in Hansen *et al.*, 2000). Although there may be species clusters at the first three peaks, for unknown reasons the last peak is unoccupied by any study species (Fig. 2a); species utilizing *Eulaema* as pollinators appear to do so with 'Eufriesea morphology' (see Armbruster, 1988, 1993; Hansen *et al.*, 2000).

We expected *Dalechampia* spp., with the least structural integration, to fall farther away from the ridge, on average, than Collinsieae, but the two groups of species were actually distributed very similarly (Fig. 2). This may be because the morphology of *Collinsia* flowers, with stamens and style enclosed in a linear keel, precludes escape from the herkogamy–accuracy trade-off, as appears to have happened in some *Dalechampia* species. Instead, herkogamy as a proportion of flower length is remarkably large during much of the flower's life in most species of *Collinsia* (see fig. 4 in Armbruster *et al.*, 2002), presumably contributing to population departure from the local adaptive optimum.

The position of the population and species means near the isometric adaptive ridge in all study systems did not appear to be simply an artefact of correlated evolution driven by overall flower or pollinator size. Although there was a positive genetic correlation between stigma and stamen lengths in one population of *D. scandens*, the slope was significantly nonisometric at 0.52. In addition, the correlation between species mean GSD and GAD remained high after conditioning on GA (the size trait best predicting pollinator size). Furthermore, GA was not a very good predictor of GAD after the effect of GSD had been removed (Fig. 3), suggesting that overall blossom–pollinator isometry is not the source of the tight GAD–GSD relationship. Additional support across all three study systems comes from a consideration of the strength and slopes of the relationships between stigma–reward and anther–reward distances in comparison with other floral trait relationships. In general, the correlations and slopes of the ASDs were much closer to the expected value of 1.0, and the intercept closer to the origin, as expected, than other trait combinations.

That *Stylidium* species and populations all fall out along the crest of the adaptive ridge can also be viewed as support of the hypothesis that fitness is highest along this isometric trajectory. Indeed, the tighter relationship in *Stylidium* (with greatest structural integration, i.e. fusion of floral parts), compared with *Collinsia* (with intermediate structural integration) and *Dalechampia* (with least structural blossom integration), suggests that the adaptive ridge is indeed narrower, as expected, in the most accurate study system.

An alternative interpretation is that the tight relationship in *Stylidium* may simply reflect a mechanical/pleiotropic constraint (see Schluter, 1996): the fusion of staminate and pistillate tissues that make up the column. By this reasoning, the perfect correlation is an automatic consequence of the structural

relationship. However, this begs the question of how and why this complex structure came to be, and leads us back to the original hypothesis that it exists as a result of selection for coordinating the positions of the anthers and stigmas during the sequential male and female phases. Future phylogenetic comparative studies of the origins, losses and modifications of the column may shed light on the selective pressures involved in its evolution.

Adaptive accuracy

The flowers of the 11 species of *Stylidium* were *c.* 15 times more accurate than the flowers and blossoms of the 15 species of Collinsieae and 28 species of *Dalechampia*, respectively. As expected this pattern parallels the trend of structural and statistical integration, with *Stylidium* flowers being the most integrated structurally, and *Dalechampia* the least. *Stylidium* is also much more statistically integrated than the other two genera (Armbruster *et al.*, 2004, 2009). The relationship between integration and floral accuracy in this case is easy to interpret. The fusion of the staminate and pistillate tissues in combination with the temporal, rather than spatial, displacement of sexual functions has allowed *Stylidium* to achieve nearly perfect mean optimality (i.e. tight correspondence of where pollen is placed on and picked up from pollinators). This, in combination with high precision, leads to high floral accuracy.

The correlation between female precision and male accuracy detected in *Collinsia* could either be the result of a causal influence of female precision on male accuracy, a causal influence of male accuracy on female precision, or the two variables being similarly influenced by a third. In this system, it seems most likely that stigma traits influence stamen traits rather than vice versa, because stigma position is programmed developmentally to change with flower age (see below). This leads to imprecision in stigma position, observed particularly in the outcrossing species of *Collinsia*.

Causes of departure from the optimum

It was interesting to note that reward–anther and reward–stigma distances tended not to covary isometrically among populations within a species in *Dalechampia*, although they did so quite strongly at the level of species means. This suggests that there may be genetic constraints, such as pleiotropy, that prevent populations from diverging optimally, even though species do so. Indeed, the trajectory observed in one sample of the *D. scandens* populations was very similar to the trajectory of the genetic correlation (0.54 vs 0.67), as would be expected if pleiotropy limited response to selection to the suboptimal trajectory of the genetic regression (Hansen *et al.*, 2003a,b). The closer approach to the adaptive trajectory by species means than by population means is consistent with the idea that evolutionary response to selection at odds with the genetic trajectory takes more time than does evolution in response to

selection that is parallel to the genetic trajectory (Schluter, 1996, 2000; Hansen & Houle, 2008). This is seen in species–population comparisons because populations have less time to diverge than do species; species usually represent more complete isolation and deeper phylogenetic branches than do populations. The differential ‘behaviour’ of populations and species is also consistent with the idea that adaptive ridge tracking requires disruption of the genetic architecture of populations, which may (or may not) be associated with speciation (Gould, 2002).

Shallower allometric slopes at lower levels of nested hierarchies (e.g. populations nested within species) have been noted in previous studies. Differences in slopes have been suggested to be statistical artefacts related to measurement error (Pagel & Harvey, 1988), but others have shown that biological explanations are much more likely (Lande, 1979; Burt, 1989; Riska, 1989, 1991; Armbruster, 1991; Hansen *et al.*, 2008). Further evidence against the artefact problem in the present study is that the slopes in question are nearly isometric, and isometric slopes do not generate the artefact (Pagel & Harvey, 1988).

Alternative resolutions of conflicting selective pressures? Selection for increased outcrossing in self-compatible species with simultaneous or overlapping sexual functions will generally favour herkogamy. This pressure will often select directly against optimality of the mean, because correspondence of the reward–anther and reward–stigma distances often results in anthers and stigmas being close together (increasing the likelihood of self-pollination). This is almost certainly the reason why *Collinsieae* have rather low accuracy for such an integrated flower [with connate petals and adnate (epipetalous) stamens]. Apparent conflict in selective pressures seems to be especially strong in the outcrossing species of *Collinsia*, which show strong herkogamy (and low optimality) over a large portion of the life of a flower, with reduced herkogamy only towards the end of flower life (Kalisz *et al.*, 1999; Armbruster *et al.*, 2002). This developmental pattern is reflected in the strong contribution of departure from the optimum and female imprecision to inaccuracy (the change in style length over the life of the flower contributes to plant- and population-level imprecision; Fig. 4b).

Low accuracy as a result of herkogamy is also probably a factor in the relatively low accuracy of many *Dalechampia* blossoms. This relationship is complicated, however, by the three-dimensional ‘escape’ from the trade-off between herkogamy and anther–stigma optimality in some species (see below). *Stylidium* spp. do not face this conflict in selection because they have escaped from the conflict through dichogamy (see below and Armbruster *et al.*, 1994, 2004).

There are thus at least two possible routes of escape from the trade-off between accuracy and herkogamy (whilst maintaining outcrossing). One is illustrated by *Stylidium*: escape in time by separating sexual functions temporally (dichogamy). Flowers initially dispense pollen for a couple of days and then

subsequently collect pollen from pollinators (Armbruster *et al.*, 1994). This escape from the trade-off may, in part, explain the much higher optimality and accuracy in *Stylidium* compared with *Dalechampia* and *Collinsia*, which are self-compatible and incompletely dichogamous.

Some populations of *D. scandens* and some species of *Dalechampia* appear, however, to escape from the trade-off by using higher dimensional space. Rather than having the reward, stigmas and anthers in a single line or plane, such that the distances are nearly additive (for example, $ASD = GSD - GAD$), as is the case for many populations and species, some species have the styles and staminate flowers diverging from the gland in a different plane or linear dimension. This ‘solution’ appears to have been employed by a number of *Dalechampia* species and evolved at least five times (Fig. 10). This system of herkogamy actually only works well because of partial dichogamy, however. In the pistillate phase (stigmas receptive, no male flowers open), reward-collecting bees contact the stigmas and transfer allogamous pollen. In the bisexual phase (stigmas receptive, one to several male flowers open), bees are much less likely to touch the stigmas, because the male flowers now form a new platform (on a different plane) on which the bees are perched. Partial rather than full dichogamy remains advantageous, as it provides the possibility of fail-safe selfing in the absence of pollinators at the end of the receptive period (reproductive assurance).

Species that are facultative or obligate selfers (for whatever reason) are not subject to the selective conflict between outcrossing (herkogamy) and the accurate fit of anthers and stigmas. Thus, we might expect them to show higher accuracy. However, this expectation is complicated by the fact that selfers may be under much more relaxed selection for accuracy, although the covariance of stamen–pistil may still be maintained to promote self-pollination (see Anderson & Busch, 2006). There was no detectable trend in *Dalechampia* for facultatively selfing species to have smaller mean departures from the optima. In *Collinsieae*, however, facultative selfing populations and species showed much lower departures from their optima than did the facultative outcrossers (Fig. 7).

Although dichogamy is common among flowering plants (Faegri & van der Pijl, 1979), it is usually interpreted as an adaptation promoting outcrossing. Although this is certainly the case, one wonders whether the ‘choice’ of dichogamy over herkogamy as a promoter might sometimes be driven by selection for anther–stigma accuracy. Future comparative studies could address this question by looking at evolutionary transitions between the two types of outcrossing promoter.

Although the use of higher dimensional space as a way to break out of the herkogamy–accuracy trade-off has not, to our knowledge, been described previously, we expect there to be many examples besides *Dalechampia*. Open chamber and ‘platform’ flowers are good candidates. Consider *Passiflora*, for example. The upright, platform flowers can achieve herkogamy (spatial separation of the anthers and stigmas) in horizontal

space by having the three stigmas positioned between the five anthers, with room to spare. However, mean optimality is determined in the vertical dimension, by the match between the corona–stigma and corona–anther distances (W. S. Armbruster, unpublished). Heterostyly is another type of escape, where having two forms of flowers and intramorph incompatibility means that optimality (reciprocity) can be high even when there is strong herkogamy (Sanchez *et al.*, 2008). Interestingly, some heterostylous species have also escaped into higher dimensional space, apparently to improve further the efficiency of intermorph pollen transfer (Armbruster *et al.*, 2006).

Concluding remarks and future research

The comparative analyses of interspecific and interpopulational data on floral accuracy supported our hypothesis that there is a fitness surface governing the interaction of reward–stigma and reward–anther distances at the species level. Indeed, conformance of species means and, to a lesser extent, population means to a positive isometric line passing through the origin strongly supports our hypothesis of an isometric adaptive ridge governing the size of structures controlling where pollen is placed on pollinators and where stigmas touch pollinators to collect pollen.

In comparing the three study systems, *Dalechampia*, *Collinsia*, and *Stylidium*, we observed considerable variation in degrees of accuracy (closeness of individuals and population means to the hypothesized adaptive ridge), with marked variation in the relative importance of phenotypic precision vs mean optimality in generating floral inaccuracy. It appears that genetic constraints on precision, as manifested through varying degrees of floral integration, impose important limits on *Dalechampia* accuracy. Inaccuracy in *Collinsia* appears to be largely a product of conflicting selective pressure promoting herkogamy (spatial separation of anthers and stigmas) during most of the life of the flower, which, in the context of the linear arrangement of fertile parts, results in low accuracy. *Stylidium* achieves high accuracy as a result of escaping the need for herkogamy by being dichogamous (temporal separation of sexual functions) and by virtue of the extreme integration of floral parts, notably the fusion of staminal and pistil tissues into a motile column.

We recommend that future investigations consider in more detail the shape of the adaptive surface controlling the coordinated evolution of the positions of pollen placement and pickup. Although we may be correct in invoking an adaptive ridge that runs along an isometric diagonal, we lack detailed insights into the shape of the ridge. Under which conditions is it broad, and under which is it narrow? This is important because, if the ridge is broad, the adaptive cost of small deviations from the optimum will be small and selection weak. If the ridge is narrow, the cost will be large and selection strong. Is the top of the ridge smooth or bumpy? Bumpiness of the

adaptive ridge seems likely to be the rule because the distributions of pollinator size and/or behaviour are usually discontinuous, as noted above.

It should be possible to gain further insight into the shape of the adaptive surface by examining the variances of the optima. For example, large optimum variances would suggest broader ridges with more gradual approach planes. It remains to be determined how to deal with this issue mathematically, however. For example, if broad ridges are associated with large variance in the optima, then maybe the optimum variance should actually be subtracted from the inaccuracy estimate rather than added to it (but cf. Armbruster *et al.*, 2009). Alternatively, the variance in the optimum could be used in a separate explicit step of mapping relative fitness onto accuracy. Clearly, there are opportunities for further theoretical and empirical development.

Acknowledgements

We thank Mark Rausher and two anonymous reviewers for comments on an earlier draft, P. H. Olsen for the photographs in Fig. 2a and the Norwegian Research Council, the Nansen Fund and the US National Science Foundation for support (grants DEB-9318640, DEB-9708333, DEB-0324808 and DEB-0444745 to W.S.A., and DEB-0444157 to T.F.H.).

References

- Anderson IA, Busch JW. 2006. Relaxed pollinator-mediated selection weakens floral integration in self-compatible taxa of *Leavenworthia* (Brassicaceae). *American Journal of Botany* **93**: 860–867.
- Armbruster WS. 1984. The role of resin in angiosperm pollination: ecological and chemical considerations. *American Journal of Botany* **71**: 1149–1160.
- Armbruster WS. 1988. Multilevel comparative analysis of morphology, function, and evolution of *Dalechampia* blossoms. *Ecology* **69**: 1746–1761.
- Armbruster WS. 1990. Estimating and testing the shapes of adaptive surfaces: the morphology and pollination of *Dalechampia* blossoms. *American Naturalist* **135**: 14–31.
- Armbruster WS. 1991. Multilevel analyses of morphometric data from natural plant populations: insights into ontogenetic, genetic, and selective correlations in *Dalechampia scandens*. *Evolution* **45**: 1229–1244.
- Armbruster WS. 1993. Evolution of plant pollination systems: hypotheses and tests with the neotropical vine *Dalechampia*. *Evolution* **47**: 1480–1505.
- Armbruster WS. 1996. Evolution of floral morphology and function: an integrated approach to adaptation, constraint, and compromise in *Dalechampia* (Euphorbiaceae). In: Lloyd D, Barrett SCH, eds. *Floral biology*. New York, NY, USA: Chapman & Hall, 241–272.
- Armbruster WS. 2001. Evolution of floral form: electrostatic forces, pollination, and adaptive compromise. *New Phytologist* **152**: 181–183.
- Armbruster WS. 2002. Can indirect selection and genetic context contribute to trait diversification? A transition-probability study of blossom-colour evolution in two genera. *Journal of Evolutionary Biology* **15**: 468–486.
- Armbruster WS, Baldwin BG. 1998. Switch from specialized to generalized pollination. *Nature* **394**: 632–632.
- Armbruster WS, Edwards ME, Debevec EM. 1994. Character displacement generates assemblage structure of Western Australian triggerplants (*Stylidium*). *Ecology* **75**: 315–329.

- Armbruster WS, Hansen TF, Pélabon C, Pérez-Barrales R, Maad J. 2009. The adaptive accuracy of flowers: measurement and microevolutionary patterns. *Annals of Botany* 103: 1529–1545.
- Armbruster WS, Martin P, Kidd J, Stafford R, Rogers DG. 1995. Reproductive significance of indirect pollen-tube growth in *Dalechampia* (Euphorbiaceae). *American Journal of Botany* 82: 51–56.
- Armbruster WS, Mulder CPH, Baldwin BG, Kalisz S, Wessa B, Nute H. 2002. Comparative analysis of late floral development and mating-system evolution in tribe Collinsieae (Scrophulariaceae, s.l.). *American Journal of Botany* 89: 37–49.
- Armbruster WS, Pélabon C, Hansen TF, Mulder CPH. 2004. Floral integration, modularity, and precision: distinguishing complex adaptations from genetic constraints. In: Pigliucci M, Preston K, eds. *Phenotypic integration: studying the ecology and evolution of complex phenotypes*. Oxford, UK: Oxford University Press, 23–49.
- Armbruster WS, Pérez-Barrales R, Arroyo J, Edwards ME, Vargas P. 2006. Three-dimensional reciprocity of floral morphs in wild flax (*Linum suffruticosum*): a new twist on heterostyly. *New Phytologist* 171: 581–590.
- Armbruster WS, Schwaegerle KE. 1996. Causes of covariation of phenotypic traits among populations. *Journal of Evolutionary Biology* 9: 261–276.
- Arnold SJ, Bürger R, Hohenlohe PA, Ajie BC, Jones AG. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* 62: 2451–2461.
- Arnold SJ, Pfrender ME, Jones AG. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112/113: 9–32.
- Bell G. 1989. A comparative method. *American Naturalist* 133: 553–571.
- Bradshaw AD. 1991. Genostasis and the limits to evolution. *Philosophical Transactions of the Royal Society of London B* 333: 289–305.
- Burt A. 1989. Comparative methods using phylogenetically independent contrasts. In: Harvey PA, Partridge L, eds. *Oxford surveys in evolutionary biology*, vol. 6. Oxford, UK: Oxford University Press, 33–53.
- Conner J, Via S. 1993. Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. *Evolution* 47: 704–711.
- Conner JK. 1997. Floral evolution in wild radish: the roles of pollinators, natural selection, and genetic correlations among traits. *International Journal of Plant Sciences* 158: S108–120.
- Conner JK. 2002. Genetic mechanisms of floral trait correlations in a natural population. *Nature* 420: 407–410.
- Darwin C. 1877. *The various contrivances by which orchids are fertilised by insects*, 2nd edn. London, UK: Murray.
- Endler JA. 1986. *Natural selection in the wild*. Princeton, NJ, USA: Princeton University Press.
- Endler JA. 1995. Multiple-trait coevolution and environmental gradients in guppies. *Trends in Ecology and Evolution* 10: 22–29.
- Fægri K, van der Pijl L. 1979. *The principles of pollination ecology*, 3rd edn. Oxford, UK: Pergamon, 244.
- Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.
- Gavrilets S. 2004. *Fitness landscapes and the origin of species*. Princeton, NJ, USA: Princeton University Press.
- Gould SJ. 2002. *The structure of evolutionary theory*. Cambridge, MA, USA: Belknap/Harvard Press.
- Grant V. 1971. *Plant speciation*. New York, NY, USA: Columbia University Press.
- Hansen TF, Armbruster WS, Antonsen L. 2000. Comparative analysis of character displacement and spatial adaptations as illustrated by the evolution of *Dalechampia* blossoms. *American Naturalist* 156: S17–S34.
- Hansen TF, Armbruster WS, Carlson ML, Pélabon C. 2003a. Evolvability and genetic constraint in *Dalechampia* blossoms: Genetic correlations and conditional evolvability. *Journal of Experimental Zoology. Molecular and Developmental Evolution* 296B: 23–39.
- Hansen TF, Carter AJR, Pélabon C. 2006. On adaptive accuracy and precision in natural populations. *American Naturalist* 168: 168–181.
- Hansen TF, Houle D. 2004. Evolvability, stabilizing selection, and the problem of stasis. In: Pigliucci M, Preston K, eds. *Phenotypic integration: studying the ecology and evolution of complex phenotypes*. Oxford, UK: Oxford University Press, 130–150.
- Hansen TF, Houle D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* 21: 1201–1219.
- Hansen TF, Pélabon C, Armbruster WS, Carlson ML. 2003b. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. *Journal of Evolutionary Biology* 16: 754–766.
- Hansen TF, Pienaar J, Orzack SH. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62: 1965–1977.
- Harder LD, Wilson WG. 1998. Theoretical consequences of heterogeneous transport conditions for pollen dispersal by animals *Ecology* 79: 2789–2807.
- Heinrich B, Raven PH. 1972. Energetics and pollination ecology. *Science* 176: 597–602.
- Johnson SD, Steiner KE. 1997. Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (Orchidaceae). *Evolution* 51: 45–53.
- Jones AG, Arnold SJ, Bürger R. 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* 57: 1747–1760.
- Jones AG, Arnold SJ, Bürger R. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* 58: 1639–1654.
- Kalisz S, Vogler D, Fails B, Finer M, Shepard E, Herman T, Gonzales R. 1999. The mechanism of delayed selfing in *Collinsia verna* (Scrophulariaceae). *American Journal of Botany* 86: 1239–1247.
- Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. *Evolution* 33: 402–416.
- Lande R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* 94: 203–215.
- Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Lankinen A, Skogsmyr I. 2001. Evolution of pistil length as a choice mechanism for pollen quality. *Oikos* 92: 81–90.
- Li CC. 1975. *Path analysis – a primer*. Pacific Grove, CA, USA: Boxwood Press.
- Maad J. 2000. Phenotypic selection in hawkmoth-pollinated *Platanthera bifolia*: targets and fitness surfaces. *Evolution* 54: 112–123.
- Maad J, Alexandersson R. 2004. Variable selection in *Platanthera bifolia* (Orchidaceae): phenotypic selection differed between sex functions in a drought year. *Journal of Evolutionary Biology* 17: 642–650.
- Mulcahy D. 1983. Models of pollen tube competition in *Geranium maculatum*. In: Real L, ed. *Pollination biology*. New York, NY, USA: Academic Press, 151–161.
- Nesse RM. 2005. Maladaptation and natural selection. *Quarterly Review of Biology* 80: 62–70.
- Nilsson LA. 1988. The evolution of flowers with deep corolla tubes. *Nature* 334: 147.
- O'Connell LM, Johnston MO. 1998. Male and female pollination success in a deceptive orchid, a selection study. *Ecology* 79: 1246–1260.
- Orzack SH, Sober E. 1994a. Optimality models and the test of adaptationism. *American Naturalist* 143: 361–380.
- Orzack SH, Sober E. 1994b. How (not) to test an optimality model. *Trends in Ecology and Evolution* 9: 265–267.
- Pagel MD, Harvey PH. 1988. The taxon-level problem in the evolution of mammalian brain size: fact and artefacts. *American Naturalist* 132: 344–359.
- Pélabon C, Hansen TF. 2008. On the adaptive accuracy of directional asymmetry in insect wing size. *Evolution* 62: 2855–2867.

- Polly D. 2008. Developmental dynamics and G-matrices: can morphometric spaces be used to model phenotypic evolution? *Evolutionary Biology* 35: 83–96.
- Purvis A, Rambaut A. 1995. Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Computer Applications in Biological Sciences* 11: 247–251.
- Revell LJ. 2007. The G matrix under fluctuating correlational mutation and selection. *Evolution* 61: 1857–1872.
- Riska B. 1989. Composite traits, selection response, and evolution. *Evolution* 43: 1172–1191.
- Riska B. 1991. Regression models in evolutionary allometry. *American Naturalist* 138: 283–299.
- Sanchez JM, Ferrero V, Navarro L. 2008. A new approach to the quantification of degree of reciprocity in distylous (sensu lato) plant populations. *Annals of Botany* 102: 463–472.
- Schluter D. 1988. Estimating the form of natural-selection on a quantitative trait. *Evolution* 42: 849–861.
- Schluter D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50: 1766–1774.
- Schluter D. 2000. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.
- Schluter D, Nychka D. 1994. Exploring fitness surfaces. *American Naturalist* 143: 597–616.
- Schluter D, Price TD, Rowe L. 1991. Conflicting selection pressures and life-history trade-offs. *Proceedings of the Royal Society of London Series B, Biological Sciences* 246: 11–17.
- Simpson GG. 1944. *Tempo and mode in evolution*. New York, USA: Columbia University Press.
- Sokal RR, Rohlf FJ. 1981. *Biometry*. San Francisco, CA, USA: Freeman.
- Stebbins GL. 1951. Natural selection and the differentiation of angiosperm families. *Evolution* 5: 299–234.
- Stebbins GL. 1974. *Flowering plants. Evolution above the species level*. Cambridge, MA, USA: Harvard University Press, 397.
- Steiner KE, Whitehead VB. 1990. Pollinator adaptation to oil-secreting flowers – *Rediviva* and *Diascia*. *Evolution* 44: 1701–1707.
- Steiner KE, Whitehead VB. 1991. Oil flowers and oil bees: further evidence for pollinator adaptation. *Evolution* 45: 1493–1501.
- Strauss SY, Irwin RE. 2004. Ecological and evolutionary consequences of multispecies plant–animal interactions. *Annual Review of Ecology, Evolution, and Systematics* 35: 435–466.
- Thompson JN, Nuismer SL, Gomulkiewicz R. 2002. Coevolution and maladaptation. *Integrative and Comparative Biology* 42: 381–387.
- Turelli M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution* 42: 1342–1347.
- Wallace B. 1975. Hard and soft selection revisited. *Evolution* 29: 465–473.
- Webster GL, Webster BD. 1972. Morphology and relationships of *Dalechampia scandens* (Euphorbiaceae). *American Journal of Botany* 59: 573–586.
- Williams GC. 1992. *Natural selection. Domains, levels, and challenges*. Oxford, UK: Oxford University Press.
- Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

About New Phytologist

- New Phytologist is owned by a non-profit-making **charitable trust** dedicated to the promotion of plant science, facilitating projects from symposia to open access for our Tansley reviews. Complete information is available at www.newphytologist.org.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as-ready' via *Early View* – our average submission to decision time is just 29 days. Online-only colour is **free**, and essential print colour costs will be met if necessary. We also provide 25 offprints as well as a PDF for each article.
- For online summaries and ToC alerts, go to the website and click on 'Journal online'. You can take out a **personal subscription** to the journal for a fraction of the institutional price. Rates start at £139 in Europe/\$259 in the USA & Canada for the online edition (click on 'Subscribe' at the website).
- If you have any questions, do get in touch with Central Office (newphytol@lancaster.ac.uk; tel +44 1524 594691) or, for a local contact in North America, the US Office (newphytol@ornl.gov; tel +1 865 576 5261).