ANNALS OF
BOTANY
Founded 1887

VIEWPOINT

Using phenotypic manipulations to study multivariate selection of floral trait associations

Diane R. Campbell*

Department of Ecology & Evolutionary Biology, University of California, Irvine, CA 92697, USA

• *Background* A basic theme in the study of plant–pollinator interactions is that pollinators select not just for single floral traits, but for associations of traits. Responses of pollinators to sets of traits are inherent in the idea of pollinator syndromes. In its most extreme form, selection on a suite of traits can take the form of correlational selection, in which a response to one trait depends on the value of another, thereby favouring floral integration. Despite the importance of selection for combinations of traits in the evolution of flowers, evidence is relatively sparse and relies mostly on observational approaches.

• *Scope* Here, methods for measuring selection on multivariate suites of floral traits are presented, and the studies to date are reviewed. It is argued that phenotypic manipulations present a powerful, but rarely used, approach to teasing apart the separate and combined effects of particular traits. The approach is illustrated with data from studies of alpine plants in Colorado and New Zealand, and recommendations are made about several features of the design of such experiments.

• *Conclusions* Phenotypic manipulations of two or more traits in combination provide a direct way of testing for selection of floral trait associations. Such experiments will be particularly valuable if rooted in hypotheses about differences between types of pollinators and tied to a proposed evolutionary history.

Key words: Colour, correlational selection, experiment, floral integration, multivariate selection, phenotypic manipulation, pollination syndrome, pollinator visitation.

INTRODUCTION

Flowers exhibit an astonishing variety of colours, shapes, sizes and positioning of reproductive parts. There is a long tradition of attempting to explain this variety in terms of suites of traits that are associated with one another and/or with a particular type of pollinator. These ideas have taken several forms. One of the best known is the notion of pollination syndromes, in which flowers are grouped into categories that are associated with particular pollinator types (Faegri and van der Pijl, 1966). In such a scheme, flowers of a particular style share not one trait, but a suite of traits. For example, in the Baker and Hurd (1968) scheme, the sphingophily (hawkmoth) syndrome is characterized by flowers that are white or pale, with a strong, usually sweet, odour, usually actinomorphic and held horizontal or pendent, with a deep narrow tube to the corolla, and ample and concealed nectar. Such a suite of characters is thought to reflect convergent evolution of unrelated species due to selection imposed by the perceptual abilities of a pollinator and its ability to transfer pollen efficiently, although the evidence for this process and its applicability is controversial (Waser *et al.*, 1996; Ollerton *et al.*, 2009).

Testing the role of selection by pollinators is helped by an understanding of the myriad forms it can take. Even for a single floral trait, selection can be directional, stabilizing or disruptive. In directional selection, fitness increases with more extreme values of a trait, for example the increase in hawkmoth-mediated pollen export for flowers with an especially deep floral tube in orchids of the genus *Platanthera* (Nilsson, 1988). In stabilizing selection, an intermediate value is favoured, as illustrated by the peak in male reproductive success for plants of the noctuid moth-pollinated *Silene latifolia* (Caryophyllaceae) that have intermediate calyx diameters (Wright and Meagher, 2004). Disruptive selection is a pattern in which extreme values are favoured, as in the case of pollinator visitation by the combination of hummingbirds and hawkmoths as a function of corolla width in a hybrid zone of *Ipomopsis* (Polemoniaceae; Campbell, 2003). When selection on a whole suite of floral traits is considered, selection may also take the form of correlational selection, defined as selection on a combination of traits that is not predictable from examining selection on the two single traits, one at a time. For example, in correlational selection for deep, narrow corolla tubes, the effect of tube length on fitness would depend on the width of the tube. In the simplest case, if each of two traits were controlled by a single gene, correlational selection would be equivalent to epistasis for fitness.

These types of selection on floral traits can be quantified, by examining their separate and combined relationship with fitness. A widely used approach for quantitative traits is phenotypic selection analysis (Lande and Arnold, 1983). Consider two traits, as illustrated in Fig. 1 for corolla length and stamen length, and examine the multiple regression of relative fitness (w) on the two traits (X_1, X_2), each standardized to a mean of zero and variance of 1.

$$w = \alpha + \beta_1 X_1 + \beta_2 X_2 + 0.5\gamma_{11}X_1^2 + 0.5\gamma_{22}X_2^2$$
$$+ \gamma_{12}X_1X_2 + \varepsilon \tag{1}$$

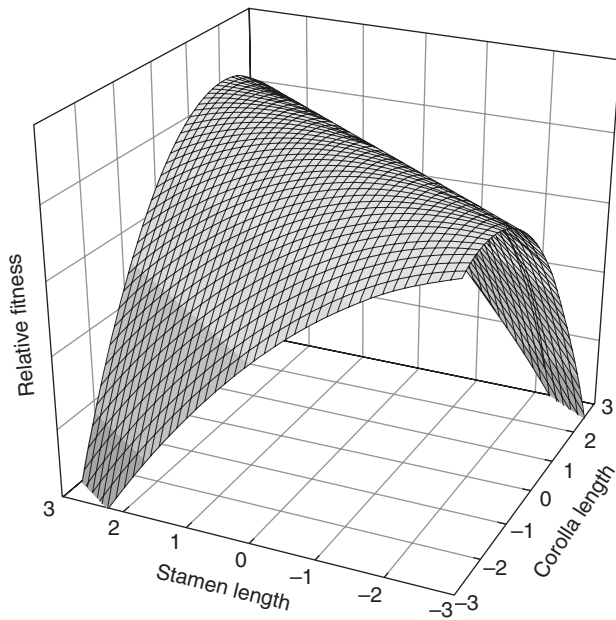* For correspondence. E-mail drcampbe@uci.edu

F$_{\text{IG}}$. 1. Hypothetical selection surface showing correlational selection on two traits: corolla length (X_1) and stamen length (X_1). The surface is based on eqn (1): $w = \alpha + \beta_1 X_1 + \beta_2 X_2 + 0{\cdot}5\gamma_{11}X_1^2 + 0{\cdot}5\gamma_{22}X_2^2 + \gamma_{12}X_1 X_2 + \varepsilon$, with α, β_1 and β_2 set equal to zero, corresponding to the absence of directional selection, $\gamma_{11} = -2$ and $\gamma_{22} = -1$, corresponding to stabilizing selection on both traits, and $\gamma_{12} = 1{\cdot}41$ corresponding to correlational selection.

In this expression,β_1 and β_2 measure the strength of directional selection on each trait separately (assuming multivariate normality prior to selection, otherwise directional selection is estimated from regression with just the linear terms), γ_{11} and γ_{22} measure the strength of stabilizing or disruptive selection on each of those traits, and γ_{12} measures the strength of correlational selection (Lande and Arnold, 1983; Phillips and Arnold, 1989; Brodie, 1992). Starting with no directional selection (β_1, $\beta_2 = 0$), stabilizing selection (γ_{11}, $\gamma_{22} < 0$) favours an intermediate value for each trait. Adding moderate correlational selection (Fig. 1) produces a ridge of high fitness such that plants with short stamens and short corollas are as fit as those with long stamens and long corollas. In this example of correlational selection, what matters is the match of the two traits. This result (which depends on the value of γ_{12}) might represent a situation in which the position of the anther relative to the opening of a corolla tube is critical for efficient transfer of pollen (Conner *et al.*, 2009).

Correlational selection is of special interest because it can be an evolutionary cause of genetic and phenotypic correlations between traits. Simulation models have illustrated the evolution of genetic covariance in response to correlational selection as defined above, provided there are also pleiotropic effects of genes on the two characters (Jones *et al.*, 2003). The correlational selection in Fig. 1 would generate a tighter genetic covariance between corolla length and stamen length. Thus correlational selection is one possible cause of phenotypic integration, defined as the pattern of functional, developmental and/or genetic correlations (Pigliucci, 2003). As this definition implies, such a quantitative genetic pattern can also be produced by common developmental regulation

of two traits (Armbruster and Schwaegerle, 1996; Bissell and Diggle, 2008). The role of pollinator-mediated selection in producing floral integration has been addressed primarily by comparing the strength of correlations across types of traits or types of plants, and testing the predictions that plants dependent on more specialized pollinators should show tighter trait integration, and plants should show tighter integration of floral traits than they do of floral with vegetative traits (Berg, 1959; Conner and Via, 1993; Armbruster *et al.*, 1999; but see Ordano *et al.*, 2008).

Many of our most fundamental ideas about floral evolution, including the concept of pollination syndromes, posit that particular pollinators favour certain combinations of floral traits, such that we must view selection on flowers as a multivariate process. To date, this idea has been approached largely with observational methods. Here it is argued that experimental methods relying on phenotypic or genetic manipulation would provide a powerful and complementary means of testing this idea.

Despite long-standing interest in the idea of associated floral traits, there are remarkably few attempts to use experimental approaches (Herrera, 2001). Experimental manipulations of traits have proven a powerful approach to demonstrating selection on individual flower traits, dating back at least to Clements and Long (1923). Examples of floral traits that have been phenotypically manipulated include flower colour, nectar reward, nectar spur length, presence of a corolla lip, positions of stigma and anthers, width of the corolla, petal size and orientation of the flower (Waser and Price, 1983; Nilsson, 1988; Cresswell and Galen, 1991; Mitchell, 1993; Wilson, 1995; Campbell *et al.*, 1996; Fulton and Hodges, 1999; Cresswell, 2000; Temeles and Rankin, 2000; Castellanos *et al.*, 2004). Almost all such studies have, however, manipulated only one trait at a time. Here methods for using trait manipulation to study multivariate selection are outlined, results from case studies are presented and future directions for studying floral associations with experimental methods are previewed.

MEASURING MULTIVARIATE SELECTION

Measuring selection of floral trait associations requires more than one trait to be considered. I discuss three approaches, differing in increasing experimental control over trait values. The first approach is purely observational, in that it uses phenotypic selection methods to measure selection on combinations of traits, as well as individual traits. This can be done using standard methods of phenotypic selection analysis (Fig. 1) or by other statistical methods for describing the selection surface (Schluter and Nychka, 1994). A drawback to this approach is that it is essentially correlational and may suggest spurious relationships that are really due to effects of traits not included in the analysis (Mitchell-Olds and Shaw, 1987), or to environmentally induced correlations between floral traits and fitness (Rausher, 1992). A second approach compares the effects of manipulating one trait (e.g. flower colour) across experiments in which plants either do, or do not, vary naturally in a second trait. Finally, to avoid spurious correlations, a third approach is to manipulate two or more traits in factorial combinations.

Method 1: observational methods for estimating phenotypic selection

This method of studying selection on a set of floral traits has been used by far the most often. Whereas many of these studies have employed multiple regression to control statistically for the effect of phenotypically or genetically correlated traits, most have attempted to estimate directional selection only, and few have allowed for the possibility of correlational selection (Kingsolver *et al.*, 2001). Because this form of selection has not been emphasized in prior reviews of pollinator-mediated selection, I briefly review such studies here.

A pioneering example that included interactions between floral traits is Armbruster's (1990) estimated selection surface for pollination as a function of gland area and gland-to-stigma distance in *Dalechampia*. Three studies of orchids have tested statistically for correlational selection. Using pollinium receipt as a measure of pollination success in *Cypripedium acaule*, O'Connell and Johnston (1998) detected an interaction between the effects of flower height and labellum length. Maad (2000) found weak correlational selection on a combination of flowering date and plant size in *Platanthera bifolia*. Also, correlational selection favoured a combination of high flower number and deep nectaries in the orchid *Cyclopogon elatus* (Benitez-Vieyra *et al.*, 2006). Cresswell and Galen (1991) observed bumble-bees at four classes of naturally varying flowers of *Polemonium viscosum*. Handling time was affected by an interaction between flower size and nectar reward, whereas number of flowers probed was influenced only by the reward. Gómez (2000) found that white-flowered plants with large displays make more seeds in *Lobularia maritima*, even though this plant is pollinated by insects thought to be highly generalized. Also, in *Erysimum mediohispanicum* there were non-additive effects of flower number and flower size on seed production (Gómez *et al.*, 2006). A full understanding of correlational selection requires not just detection of a trait interaction but demonstration of a functional mechanism to explain it. One example is provided by *Raphanus raphanistrum* (Brassicaceae), in which there is stabilizing selection for the difference between stamen and corolla tube length, which is equivalent to correlational selection on the two traits combined (Conner, 2006; Conner *et al.*, 2009). In this case, an intermediate anther position appears most effective in placing pollen onto sweat bee pollinators.

Illustration of methods using Ipomopsis. To illustrate the shape of a selection surface and compare this observational method with experimental methods for estimating selection on trait combinations, I re-analyse data on hummingbird visitation at flowers of *Ipomopsis* (Polemoniaceae), previously reported by Campbell *et al.* (1997). This study examined the rate of visitation to potted plants in experimental arrays that included plants of two species and their natural hybrids. The red-flowered *I. aggregata* also has relatively wide, short corolla tubes, in comparison with the white-flowered *I. tenuituba*. The original analysis used multiple regression to estimate directional selection on three traits: colour as measured by maximal optical density between a wavelength of 507 and 530 nm, corolla width and corolla length. I re-analysed these data to test for other forms of selection by adding quadratic

terms and two-way products of traits to the multiple regression. No selection on corolla length was detected, so I dropped that trait for simplicity. Nectar reward was not measured, and it is known to influence hummingbird visitation (Mitchell, 1993). For this and other reasons described below, the analysis should be considered an illustration of methods for measuring selection of trait associations, rather than a definitive example.

Fitting eqn (1) revealed an interaction between the effects of flower colour and corolla width ($F_{1,71} = 6.20$, $P < 0.0151$) in addition to directional selection favouring large values for both traits ($P < 0.0023$ for width and $P < 0.0105$ for colour), but no stabilizing or disruptive selection. The interaction coefficient had a negative sign and thus was in the opposite direction from predicted if hummingbirds are cuing in on the reddest, widest flowers, as might be expected from the hummingbird syndrome. Instead, increasing the width of the flower increased visit rate more than expected for the pale flowers, as compared with the red ones, such that expected visitation is actually highest for a combination of narrow and red flowers (Fig. 2A), in part due to one exceptionally attractive plant. When the data are broken down by type of plant, for hybrid plants with intermediate colours, pollinators responded by preferentially visiting those with wider-tubed flowers, but showed no such response for red-flowered plants of *I. aggregata* (which were visited heavily overall) or the very pale-coloured flowers of *I. tenuituba* (Fig. 2A).

One statistical limitation of this approach is that a combination of linear and quadratic terms is not guaranteed to describe selection well, even for a single trait. An alternative way to describe the selection surface is by measuring selection on a combination of traits, for example with projection pursuit regression (Schluter and Nychka, 1994). To my knowledge, this method has rarely been used with floral traits (Maad, 2000). It starts by finding a linear combination of the original traits that explains as much of the variation in the fitness component (here visitation rate) as possible. In this case, the first projection loads only slightly more heavily on standardized corolla width compared with flower colour. Notice that the projection does not go through the direction that explains the greatest variation in the explanatory traits, but a direction that explains variation in pollinator visitation (Fig. 2A). Along this slice through the selection surface, visit rate is predicted to rise steeply. Predicted visit rate can now be fitted with a non-parametric cubic spline regression (Fig. 2B), rather than assuming that the form must be described by a certain number of parameters. In this case, the pursuit projection is almost perfectly linear, suggesting that the selection by hummingbirds is mostly directional, favouring red flowers and wide tubes, with a weak contribution of any correlational selection.

Further analysis of pollinator responses would benefit from experimental manipulation to extend the range of trait values to create critical combinations. In this data set, for example, there are few plants with narrow, red flowers. It would also be valuable to use this method to consider the form of selection generated by the combined visitation of hummingbirds and hawkmoths in this *Ipomopsis* hybrid zone. Since the two species are not fully reproductively isolated (Campbell, 2003), it is of interest to consider selection on the full range of flowers by both pollinator types combined. Such selection could take the form of correlational selection. For example,
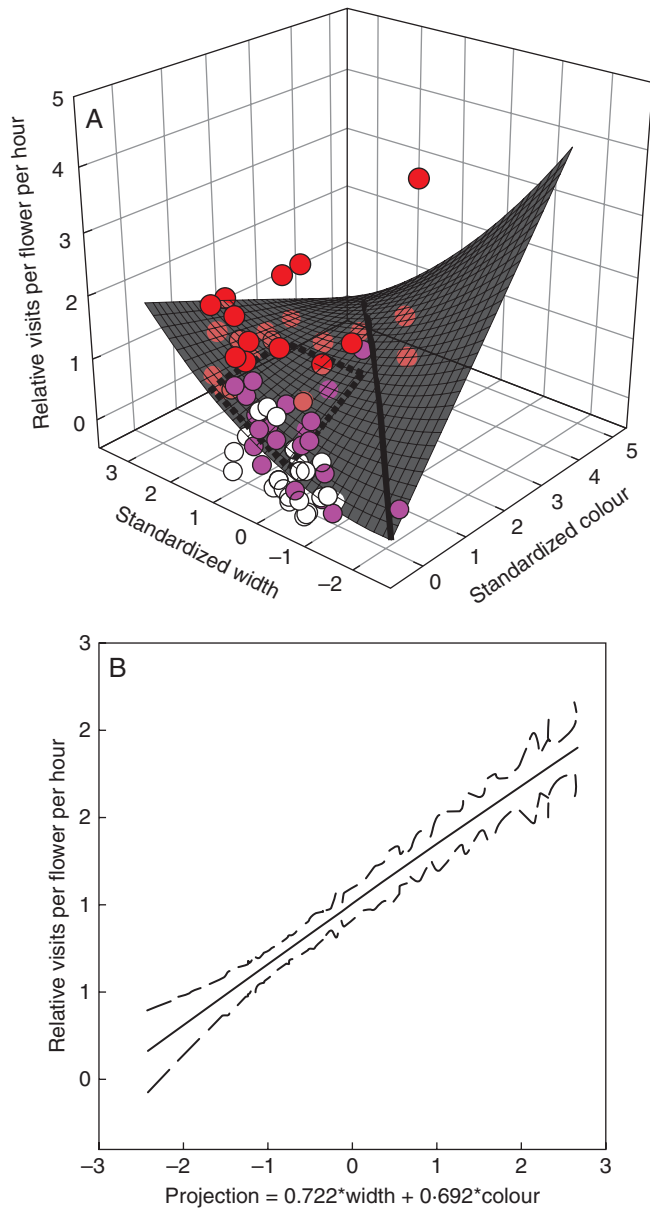
FIG. 2. Selection based on hummingbird visitation rates to plants of *Ipomopsis* in experimental arrays studied by Campbell *et al.* (1997). (A) The surface shows the best quadratic approximation to the selection surface as defined in Fig. 1, where Y = relative visit rate (absolute visit rate divided by mean visit rate), X_1 = optical density (a measure of colour) standardized to a mean of zero and variance of 1, and X_2 = corolla width standardized to a mean of zero and variance of 1. For this data set, $\gamma_{12} = -0.267$ ($P < 0.0151$). Filled circles show the data, with red indicating *I. aggregata*, pink indicating hybrid and white indicating *I. tenuituba*. The dashed lines connect the predicted visitation for plants that have the four possible combinations of mean trait values for the two species. The solid line shows the direction of the first projection obtained using Version 1·3 (2003) of the projection pursuit regression program of Schluter and Nychka available at http://www.zoology.ubc.ca/~schluter/software.html. (B) The cubic spline estimate of visitation as a function of the first projection. The solid line shows the predicted value, and the dotted lines show ± 1 s.e. Both coefficients in the projection are significantly different from zero based on 1000 bootstrap values. After pilot runs, the smoothing parameter λ was set to 4 to minimize generalized cross-validation function (GCV) at 0·234.

if hummingbirds strongly prefer short, wide tubes while hawk-moths prefer long, narrow tubes, then the effect of tube length on overall visitation would depend on the width of the tube. Even if each pollinator were exerting directional selection on the two traits, the combined effect of two pollinators could be the production of two peaks on a selection surface, one corresponding to bird visitation and one to moth visitation, with a fitness minimum in between.

Method 2: manipulation of one trait with natural variation in other traits

A second method of investigating selection of trait associations involves manipulating one trait against a background in which other traits vary to different degrees. This approach goes beyond a simple manipulation of a single trait in that it allows assessment of whether the effect of that trait is context dependent. My colleagues and I have used this approach with experimental arrays of *I. aggregata*, *I. tenuituba* and natural hybrids, in this case to analyse the effects of flower colour vs. all other traits that separate the species (Meléndez-Ackerman and Campbell, 1998).

Although rarely included in studies of multivariate selection, flower colour is a good choice of a trait for experimental methods, as it is relatively easily manipulated. Flowers can be painted to match certain reflectance spectra (Meléndez-Ackerman *et al.*, 1997), and control flowers painted their natural colour can serve as controls to rule out the effects of the paint itself. Even so, a recent review by Rausher (2008) found just two cases where flower colour had been phenotypically manipulated: *Delphinium nelsonii* and *Ipomopsis*, both studied in sub-alpine meadows in Colorado. In *D. nelsonii*, the colour contrast provided by a nectar guide reduced handling time and increased visitation by hummingbirds and bumble-bees (Waser and Price, 1983, 1985).

In the *Ipomopsis* study, three types of arrays were observed: a single-species array with flowers painted different colours, a mixed array with all natural variation intact, and a mixed array with flowers painted the same colour but with other natural variation present (Meléndez-Ackerman and Campbell, 1998). The three types of arrays were used to sort out the role of colour relative to the roles of other floral traits in influencing visitation. A single-species array with flower colour manipulated tests for the role of colour alone. A mixed array with all plant types painted the same colour tests for the role of other floral traits. An unmanipulated array tests for the combined influence of colour and other floral traits that differ between the species. Ideally, arrays would be observed simultaneously to assess the relative effects of each combination without confounding changes in pollinator behaviour over time. In our study, we found by far the largest difference in fitness (seeds set and seeds sired) in the array with all natural variation intact, which, in combination with visitation data, suggested that the hummingbird pollinators responded to a combination of traits and not just colour (Meléndez-Ackerman and Campbell, 1998). The differences

in visitation with all traits intact was larger (relative fitnesses of *I. aggregata*, hybrids and *I. tenuituba* = 1·00, 0·25 and 0·27, respectively) than predicted by adding the fitness effects of just colour (fitnesses = 1·00, 0·80 and 0·25, respectively) and the remaining traits that separate the species (fitnesses = 0·83, 1·00 and 0·96, respectively). Thus, the hummingbirds responded to traits in a non-additive way, consistent with correlational selection, although with this approach it is not possible to identify which specific traits other than colour were involved.

Floral traits in the New Zealand alpine. Method 2 as described above allows separation of the effects of flower colour from the effects of all other trait differences between two or more types of plants. The use of single flowers or flower heads instead of whole plants also allows one to dissect out effects of floral display when examining responses of flower visitors. My colleagues and I are taking this approach with flowers visited by insects in the New Zealand alpine [D. Campbell, M. Bischoff (U. Otago, New Zealand) and A. Robertson (Massey University, New Zealand), unpubl. res.]. This habitat provides an extreme test case for the role of flower colour, as about 70 % of the species in the flora have white flowers, one of the highest percentages anywhere on Earth (Wardle, 1978), even though many have congeners elsewhere that are brightly coloured (Lloyd, 1985). The preponderance of white has traditionally been attributed to the historical absence of long-tongued social bees (Gibbs, 2006), with the implication that the insects which are present do not have preferences based on colour.

The study site is a set of natural meadows in the Remarkables Mountains on the South Island of New Zealand. We have begun exploring selection on flower colour by providing choices between white and yellow flowers (the second most common colour in the area). The set of experiments allows investigation of whether responses of particular insect pollinators to colour depend on the context of other floral traits. Each experimental array contained 16 cut flowers, of two types, in 1·5 mL tubes filled with water and spaced 10 cm apart. We calculated the proportion of visits made in a foraging bout to one type of flower and compared its mean with the null hypothesis of 0·5 using a one-sample *t*-test, following arcsin transformation.

At unmanipulated arrays containing eight flowers of each of two species, both colletid bees (*Hylaeus* sp.) and syrphid flies (*Allograpta* spp.) showed preferences for certain flower species. We next performed manipulations of colour in single-species arrays of *Brachyscome sinclarii* and *Brachyglottis bellidioides* (both Asteraceae). Acrylic paints were used to match the visible spectra of white flowers to those of naturally yellow flowers, and vice versa. These paints also block UV reflectance. Since the experiments involved controls of flowers painted both colours, UV reflectance was absent in both treatments, and responses apply only to signals in the human-visible range. These manipulations resulted in overvisitation of yellow flowers compared with white flowers for both syrphid flies and colletid bees when visiting *Brachyglottis bellidioides*, but not for syrphid flies visiting *B. sinclarii*, which is normally white (Fig. 3A).

Both types of insects also showed preferences based on other floral traits, and so it is of interest to ask how the two kinds of traits interact. For *Allograpta*, responses to other traits depended on flower colour, with the flies switching preference from *Brachyscome* to *Brachyglottis* when flowers of both species were painted yellow rather than white (Fig. 3B). Furthermore, they showed a preference for yellow only when foraging on the normally yellow-flowered species, *B. bellidioides*. The interaction between effects of the two types of traits can also be seen by comparing preferences in the three types of arrays. With all flower differences intact, these syrphid flies made 67 % of visits to *Brachyglottis*. With only colour varying (Fig. 3A), the average of the two responses (across the two species) was 62 % of visits to the yellow flowers, which is consistent with their preference for *Brachyglottis*. However, when flowers of the two species were painted the same colour, so that only morphology and rewards differed, they actually preferred on average to visit *Brachyscome* instead, with the two responses averaging only 0·375 of the visits to *Brachyglottis* (Fig. 3B). Creation of other trait combinations is still required to understand fully how these insects respond to trait associations.

Method 3: factorial manipulation of traits

The most direct way to test for selection of trait associations is by producing experimental combinations of traits in a factorial design, through either phenotypic or genetic manipulation. This method allows testing directly for (*a*) responses of pollinators to a suite of multiple traits, as suggested by pollination syndromes, and (*b*) responses that are not additive, i.e. for the presence of correlational selection. Surprisingly, this method has hardly been used. A few studies have examined the effects on pollination of manipulating several traits in a single flower species one at a time, for example Wilson's (1995) study of pollen removal and deposition in *Impatiens pallida*; However, it is extremely rare for multiple traits to be altered in a factorial combination. In one exception, Herrera (2001) reduced or completely excised the upper and/or lower corolla lip of *Lavandula latifolia*. The normal, non-manipulated phenotype had no pollination or fecundity advantage, arguing against the presence of correlational selection, but it is also a case in which no effect of either trait by itself was detected either. To my knowledge, no other studies of selection using phenotypic manipulation of two or more floral traits in factorial combination have been reported. Recently, however, Kessler *et al.* (2008) used genetic manipulation to block expression of a floral fragrance attractant and a nectar repellant in all combinations in *Nicotiana attenuata*. Both the repellant and the attractant were required for maximal fitness. Capsule production was very low in all of the transformed plants compared with the wild type, and correlational selection for a combination of scents was supported.

DESIGN OF FUTURE EXPERIMENTS

The basic idea of manipulating traits in combination to assess selection of floral trait associations seems straightforward. However, many features of the experimental design will require careful attention.
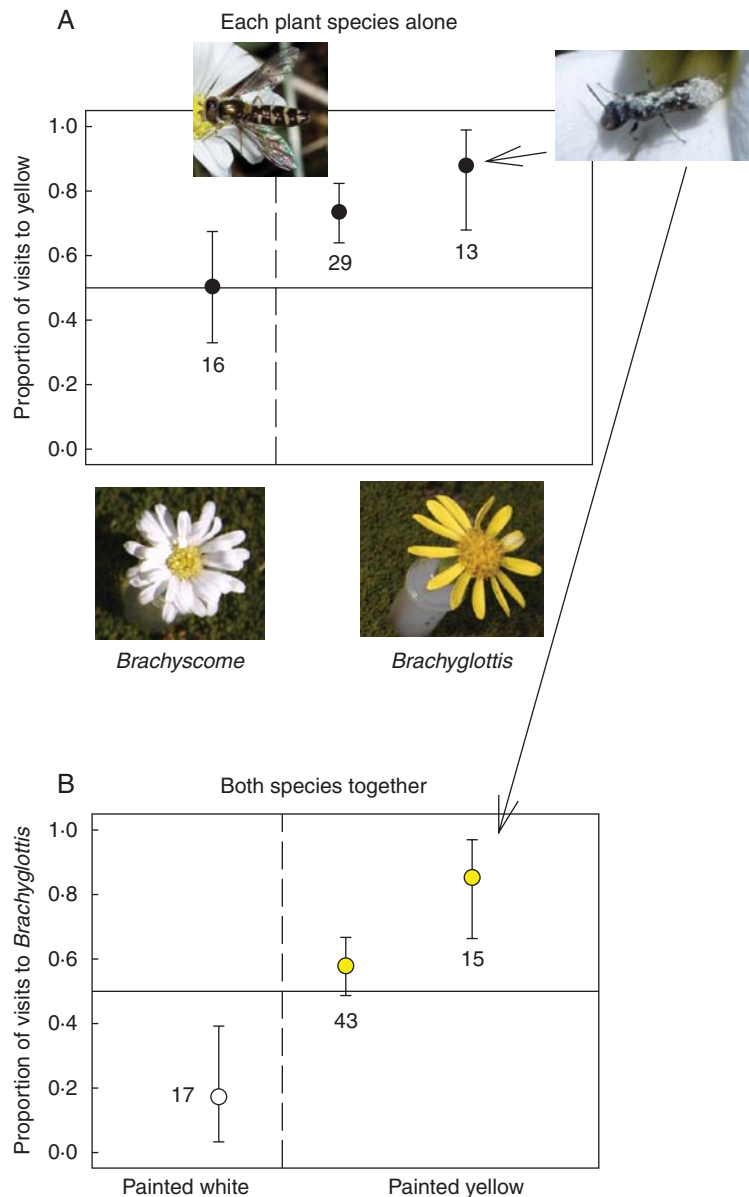
FIG. 3. Proportion of visits made to traits characteristic of *Brachyglottis bellidioide*s rather than *Brachyscome sinclarii* for insects in the New Zealand alpine. Means and 95 % confidence intervals are shown for colletid bees (*Hylaeus*) on the right as indicated by the arrows and for syrphid flies (*Allograpta*) in all other cases. Sample sizes are numbers of foraging bouts. The solid line represents the null hypothesis. (A) Proportion of visits to yellow flowers in a two-way choice between flowers painted white (to resemble *Brachyscome*) vs. yellow (to resemble *Brachyglottis*). Left of the dotted line: arrays with just *Brachyscome*. Right of the dotted line: arrays with just *Brachyglottis*. (B) Proportion of visits to *Brachyglottis* in a two-way choice between flowers of the two species painted the same colour, either white (left of the dotted line) or yellow (right of the dotted line.)

Choice of traits

Perhaps the most basic is the choice of traits to manipulate. The paucity of experiments probably reflects in part a sense that progress on a multivariate question can be made more rapidly with observational studies that can easily include many traits simultaneously. However, judicious choice of a relatively small number of traits could be highly informative. For example, factorial manipulations of two or three traits critical to a particular pollinator syndrome could go a long ways towards testing its validity.

How many traits at a time?

A related question is the number of traits to manipulate simultaneously. Imagine three scenarios. In the first, natural flowers receive either extra nectar or none. In the second, this manipulation of nectar is crossed with manipulation of flower size (large vs. small) and with flower colour (coloured vs. white), producing a choice between eight types of flowers. In the third scenario, a choice is offered between white flowers with no extra nectar, white flowers with extra nectar and coloured flowers with extra nectar. A pollinator may respond very differently to the same manipulation of nectar in these scenarios.

In both of the first two scenarios there is no correlation between nectar reward and other traits. However, a pollinator might be more likely to respond to nectar in the first scenario because it is simpler to remember locations of rewarding plants when not faced with unfamiliar new morphologies. In contrast, in the third scenario extra nectar has a correlation with flower colour, which may cause pollinators to choose the coloured flowers, but show no greater tendency to visit white flowers with ample nectar as compared with white flowers with no nectar. Thus, pollinator response to one trait may depend on the range of choices presented.

This dependence of pollinator response to the range of choices means that different experiments may be required to study behaviour in the current ecological context (e.g. where trait correlations are already present as in the third scenario) and to study behaviour as it may have occurred under conditions when a new mutation had been introduced but new trait correlations had not yet evolved (as in the second scenario). Choosing the scenario relevant to the evolutionary stage of interest may be particularly important for the study of traits thought to be key innovations, such as the evolution of nectar spurs in *Aquilegia* (Hodges and Arnold, 1995). In such a case, the order of introduction of a new character may matter, for example pollinators may respond to addition of nectar only in flowers that have deep nectar spurs.

Number of trait values and some limitations to the approach

The difficulty in studying large numbers of trait combinations to characterize a multivariate selection surface fully is perhaps the biggest limitation on the experimental approach. A set of two values per trait is the minimal number required to test for interactions in effects of traits, but it does not give much information on the shape of the surface. However, if the intent is to understand differences between two hybridizing flower species, two values per trait can be extremely informative, as the trait values can be altered to match the species means. As an illustration, consider the two species of *Ipomopsis* studied in Fig. 2A. The dashed trapezoid connects the four combinations of flower colour and corolla width that correspond to the mean values for *I. aggregata* and *I. tenuituba*. An experiment employing these trait values, i.e. painting flowers to match one of the two species and manipulating width to match one of the two species, would be expected to detect correlational selection, as the change in width makes more difference for the pale-coloured category of flowers, changing predicted visitation from 0·59 to 1·32 compared with staying nearly constant (from 1·42 to 1·41) for the darkly coloured category.

Performing manipulations: phenotypic vs. genetic manipulation

Flower traits are in many ways especially amenable to phenotypic manipulation. Unlike animal surgery, sham controls for the operation may be quick and easy, e.g. applying a coat of paint to match the natural colour, inserting a syringe without sugar water, or cutting and refastening a flower to its original shape. It is critical to consider such sham treatments, as experimental manipulations can have unintended effects. As an alternative, molecular genetic tools in principle also offer the option of manipulating traits genetically, as illustrated above with a study of floral scents. Isogenic lines have also been used to study the impact of single genes for flower colour on pollinator visitation at *Mimulus* (Bradshaw and Schemske, 2003). One important difference between phenotypic and genetic manipulation approaches is that the former alters only the target trait, whereas the latter may alter other traits if there are pleiotropic effects of the gene. If the intent is to examine the evolutionary consequences of introduction of a new gene, the latter method may be the most appropriate. For example, the flavonoids responsible for some flower colours can influence vegetative as well as floral tissue (Rausher, 2008), and have been shown to protect against UV radiation and provide defence against herbivores and/or pathogens (Frey, 2004; Strauss and Whittall, 2006). Thus, the evolution of flower colour may respond to other selective pressures besides those mediated by pollinators, and phenotypic manipulation of just the flower may miss those. On the other hand, for the ecological question of whether a certain type of pollinator responds to flower colour and other traits in the way suggested by the idea of pollinator syndromes, phenotypic manipulation of the flower is a more direct approach. Finally, it would be possible to use artificial flowers to understand some of these mechanisms (Smith *et al.*, 1996), but at the considerable cost of lower realism.

Fitness components

The question of the response variables to include in the study is common to both observational and experimental approaches to the study of selection, so I will not discuss it in detail. Nevertheless, it is important to keep in mind that some aspects of floral integration are thought to arise from pollinator behavioural responses, necessitating inclusion of variables such as visitation rate or handling time (Temeles, 1996). Others are thought to depend on the way in which morphological fit of a pollinator affects the efficiency of pollen transfer, and thereby require tracking pollen removal and/or deposition (Castellanos *et al.*, 2004). Furthermore, if the intent is to show that pollinators lead to selection on combinations of traits, it is necessary to measure both the functional impacts on pollinators or pollen movement and estimates of fitness that include seed production and seeds sired.

Spatial scale

A variety of pollinators have been shown to respond differently to floral traits depending on aspects of spatial variation. Foraging theory suggests that an animal making a choice among simultaneously available resources should exhibit higher selectivity when choosing among rich and poor resources in distinct patches than when choosing among resources mixed within a patch, provided the animals know the locations of rich and poor resources (Mitchell, 1989). This is because the marginal cost of using a poor resource is lower when resources are mixed within a patch. There is evidence, from experiments with feeders, that some hummingbirds behave this way (Mitchell, 1989). Bumble-bees also respond to the spatial patterning of resources. In one study, the approach rate of bumble-bees was higher to nectar-rich plants when foraging in a sparse population of *Echium*
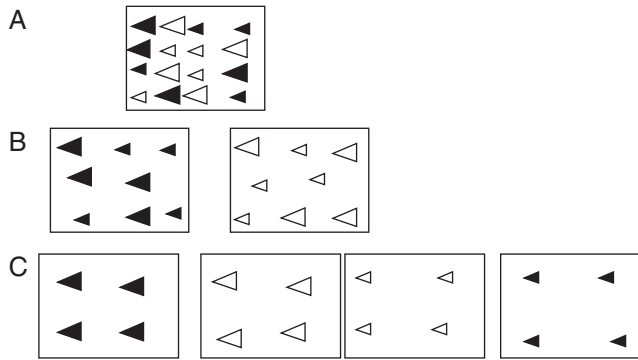
FIG. 4. Three spatial arrangements for a phenotypic manipulation experiment with two traits, flower colour (indicated by black vs. white) and flower size (large vs. small). (A) Flowers with all four factorial combinations of traits arranged in random order within a single array of plants. (B) Two arrays of different colours, each of which has flowers of both sizes. (C) Each of the four phenotypes in a separate array, with array locations chosen at random.

vulgare, but not when foraging in a dense population (Klinkhamer and van der Lugt, 2004).

This dependence of behaviour on spatial scale raises questions on how to design a study of selection on trait associations. Even with just two traits and two character states of each, an experiment could be designed in one of three ways (Fig. 4): (1) with all four flower types randomized in a single array (Fig. 4A); (2) with two arrays representing the two character states of one trait, each involving a choice between the character states of the other trait (Fig. 4B); or (3) choice among four patches of flowers, each of a different type (Fig. 4C). These arrangements may elicit different responses depending, for example, on the spatial memory of the pollinator.

Temporal scale

One of the trickiest problems in studying pollinator responses to floral traits is that many pollinators exhibit rapid learning, as illustrated by bumble-bees (Heinrich *et al.*, 1977) and hawkmoths (Riffell *et al.*, 2008). As a result, their responses to floral trait associations may depend on the frequency with which they have previously experienced them (Cresswell and Galen, 1994). Due to learning, initial responses of pollinators to novel combinations of traits presented in an experiment may not correspond to longer-term behaviour that would lead to evolutionary responses to selection. The potential effect of temporal scale is illustrated by an experiment examining responses of bumble-bees to artificial inflorescences with all four combinations of high vs. low flower number and high vs. zero sugar reward (Makino and Sakai, 2007). During the early phases of the experiment, bees chose displays with high flower number regardless of reward, but switched within 3 h to visiting primarily inflorescences that combined high flower number and reward. Thus, depending on the point in time at which results were taken, the results switched from additive effects of traits to detection of a significant interaction. Further experiments showed that the bees had learned the spatial location of the rewarding inflorescences (Makino and Sakai, 2007).

CONCLUSIONS AND SUGGESTED EXPERIMENTS

Given the long history of discussing pollination in terms of suites of associated traits, it is remarkable that there are so few experimental tests of how pollinators respond to multiple traits and whether they respond in a non-additive way to combinations of floral traits. Here I make a call for experiments testing for such responses as one critical approach to the study of pollinator-mediated evolution.

Two kinds of experiments may be especially valuable. The first type is a comparative experiment with different kinds of pollinators that tests for differences in selection hypothesized to correspond to contrasting modes of pollination. Such predictions could be based on traditional descriptions of syndromes, in which case the experiments could provide direct evidence for or against selection on the floral features thought to characterize syndromes. Furthermore, if the floral associations represent syndromes, such selection by a particular pollinator type would be expected to be similar across a broad set of plant families. Alternatevely, the predictions could be more firmly rooted in multivariate analysis of associations between floral traits and pollinators. For example, based on ordination of floral traits in *Penstemon* (Wilson *et al.*, 2004), one might predict that hummingbirds would visit *Penstemon* flowers that are red with strongly exserted anthers, whereas bees would visit primarily yellow or blue-violet flowers with inserted anthers. Finally, such predictions could be based on the perceptual and cognitive abilities of pollinators, at least in cases such as honeybees where they are well understood (Briscoe and Chittka, 2001).

A second valuable type of experiment would be one that maps on to a phylogeny in such a way that the order in which traits probably evolved can be taken into account. For example, if phylogenetic information indicates that a particular flower colour (e.g. white rather than blue) is more recently derived than a particular petal size (e.g. small rather than large), one might predict that pollinators would prefer white over blue if flowers are small (matching the step in the phylogeny), but not necessarily if flowers are large. Such information can help in deciding certain aspects of the experimental design, such as the spatial scale. The scenario just described would support the use of the spatial arrangement shown in Fig. 4B, but in which arrays of small and large flowers are observed simultaneously.

Both a proposed evolutionary history and knowledge of the sensory abilities of potential pollinators could greatly inform our study of selection and the extent to which pollinators choose particular trait combinations. Even if such information is not available, however, it is time to make use of the power of experiments, along with observational methods, to characterize the multivariate nature of pollinator responses to suites of floral traits.

LITERATURE CITED

Armbruster WS. 1990. Estimating and testing adaptive surfaces: the morphology and pollination of *Dalechampia* blossoms. *American Naturalist* **135**: 14–31.

Armbruster WS, Schwaegerle KE. 1996. Causes of covariation of phenotypic traits among populations. *Journal of Evolutionary Biology* **9**: 261–276.

Armbruster WS, Di Stilio V, Tuxill JD, Flores C, Velásquez Runk JL. 1999. Covariance and decoupling of floral and vegetative traits in nine neotropical plants: a re-evaluation of Berg's correlation–pleiades concept. *American Journal of Botany* **86**: 39–55.

Baker HG, Hurd PDJ. 1968. Intrafloral ecology. *Annual Review of Entomology* **13**: 385–414.

Benitez-Vieyra S, Medina AM, Glinos E, Cocucci AA. 2006. Pollinator-mediated selection on floral traits and size of floral display in *Cyclopogon elatus*, a sweat bee-pollinated orchid. *Functional Ecology* **20**: 948–957.

Berg RL. 1959. A general evolutionary principle underlying the origin of developmental homeostasis. *American Naturalist* **93**: 103–105.

Bissell EK, Diggle PK. 2008. Floral morphology in *Nicotiana*: architectural and temporal effects on phenotypic integration. *International Journal of Plant Sciences* **169**: 225–240.

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

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

Brodie ED. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. *Evolution* **46**: 1284–1298.

Campbell DR. 2003. Natural selection in *Ipomopsis* hybrid zones: implications for ecological speciation. *New Phytologist* **161**: 83–90.

Campbell DR, Waser NM, Price MV. 1996. Mechanisms of hummingbird-mediated selection for flower width in *Ipomopsis aggregata*. *Ecology* **77**: 1463–1472.

Campbell DR, Waser NM, Meléndez-Ackerman EJ. 1997. Analyzing pollinator-mediated selection in a plant hybrid zone: hummingbird visitation patterns on three spatial scales. *American Naturalist* **149**: 295–315.

Castellanos MC, Wilson P, Thomson JD. 2004. 'Anti-bee' and 'pro-bird' changes during the evolution of hummingbird pollination in *Penstemon* flowers. *Journal of Evolutionary Biology* **17**: 876–885.

Clements FE, Long FL. 1923. *Experimental pollination: an outline of the ecology of flowers and insects*. Washington, DC: Carnegie Institute.

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. 2006. Ecological genetics of floral evolution: In: Harder LD, Barrett SCH, eds. *Ecology and evolution of flowers*. Oxford: Oxford University Press.

Conner JK, Sahli H, Karoly K. 2009. Tests of adaptation: functional studies of pollen removal and estimates of natural selection on anther position in wild radish. *Annals of Botany* **103**: 1547–1556.

Cresswell JE. 2000. Manipulation of female architecture in flowers reveals a narrow optimum for pollen deposition. *Ecology* **81**: 3244–3249.

Cresswell JE, Galen C. 1991. Frequency dependent selection and adaptive surfaces for floral trait combinations: the pollination of *Polemonium viscosum*. *American Naturalist* **138**: 1342–1353.

Faegri K, van der Pijl L. 1966. *The principles of pollination ecology*. Oxford: Pergamon.

Frey FM. 2004. Opposing natural selection from herbivores and pathogens may maintain floral-color variation in *Claytonia virginica* (Portulacaceae). *Evolution* **58**: 2426–2437.

Fulton M, Hodges SA. 1999. Floral isolation between *Aquilegia formosa* and *Aquilegia pubescens* *Proceedings of the Royal Society B: Biological Sciences* **226**: 2247–2252.

Gibbs G. 2006. *Ghosts of Gondwana: the history of life in New Zealand*. Nelson, New Zealand: Craig Potton Publishing.

Gómez JM. 2000. Phenotypic selection and response to selection in *Lobularia maritima*: importance of direct and correlational components of natural selection. *Journal of Evolutionary Biology* **13**: 689–699.

Gómez JM, Perfetti F, Camacho JPM. 2006. Natural selection on *Erysimum mediohispanicum* flower shape: insights into the evolution of zygomorphy. *American Naturalist* **168**: 531–545.

Heinrich B, Mudge PR, Deringis PG. 1977. Laboratory analysis of flower constancy in foraging bumblebees: *Bombus ternarius* and *B. terricola*. *Behavioral Ecology and Sociobiology* **2**: 247–265.

Herrera CM. 2001. Deconstructing a floral phenotype: do pollinators select for corolla integration in *Lavandula latifolia*? *Journal of Evolutionary Biology* **14**: 574–584.

Hodges SA, Arnold ML. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society B: Biological Sciences* **262**: 343–348.

Jones AG, Arnold SJ, Burger R. 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**: 1747–1760.

Kessler D, Gase K, Baldwin IT. 2008. Field experiments with transformed plants reveal the sense of floral scents. *Science* **321**: 1200–1202.

Kingsolver JG, Hoekstra HE, Hoekstra JM, *et al.* 2001. The strength of phenotypic selection in natural populations. *American Naturalist* **157**: 245–261.

Klinkhamer PGL, van der Lugt PP. 2004. Pollinator service only depends on nectar production in sparse populations. *Oecologia (Berlin)* **140**: 491–494.

Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.

Lloyd DG. 1985. Progress in understanding the natural history of New Zealand plants. *New Zealand Journal of Botany* **23**: 707–722.

Maad J. 2000. Phenotypic selection in hawkmoth-pollinated *Platanthera bifolia*: targets and fitness surfaces. *Evolution* **54**: 112–123.

Makino TT, Sakai S. 2007. Experience changes pollinator responses to floral display size: from size-based to reward-based foraging. *Functional Ecology* **21**: 854–863.

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

Meléndez-Ackerman EJ, Campbell DR, Waser NM. 1997. Hummingbird behavior and mechanisms of selection on flower color in *Ipomopsis*. *Ecology* **78**: 2532–2541.

Mitchell RJ. 1993. Adaptive significance of *Ipomopsis aggregata* nectar production: observation and experiment in the field. *Evolution* **47**: 25–35.

Mitchell WA. 1989. Informational constraints on optimally foraging hummingbirds. *Oikos* **55**: 145–154.

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

Nilsson LA. 1988. The evolution of flowers with deep corolla tubes. *Nature* **334**: 147–149.

O'Connell LM, Johnston MO. 1998. Male and female pollination success in a deceptive orchid: a selection study. *Ecology* **79**: 1246–1260.

Ollerton J, Alarcón R, Waser NM, *et al.* 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**: 1471–1480.

Ordano M, Fornoni J, Boege K, Domínguez CA. 2008. The adaptive value of phenotypic floral integration. *New Phytologist* **179**: 1183–1192.

Phillips PC, Arnold SJ. 1989. Visualizing multivariate selection. *Evolution* **43**: 1209–1222.

Pigliucci M. 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecology Letters* **6**: 265–272.

Rausher MD. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* **46**: 616–626.

Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences* **169**: 7–21.

Riffell JA, Alarcon R, Abrell L, Davidowitz G, Bronstein JL, Hildebrand JG. 2008. Behavioral consequences of innate preferences and olfactory learning in hawkmoth–flower interactions. *Proceedings of the National Academy of Sciences, USA* **105**: 3404–3409.

Schluter D, Nychka D. 1994. Exploring fitness surfaces. *American Naturalist* **143**: 597–616.

Smith CE, Stevens JT, Temeles EJ, Ewald PW, Hebert RJ, Bonkovsky RL. 1996. Effect of floral orifice width and shape on hummingbird–flower interactions. *Oecologia (Berlin)* **106**: 482–492.

Strauss SY, Whittall JB. 2006. Non-pollinator agents of selection on floral traits. In: Harder L, Barrett Seds. *Ecology and evolution of flowers*. Oxford, UK: Oxford University Press.

Temeles EJ. 1996. A new dimension to hummingbird–flower relationships. *Oecologia (Berlin)* **105**: 517–523.

Temeles EJ, Rankin AG. 2000. Effect of the lower lip of *Monarda didyma* on pollen removal by hummingbirds. *Canadian Journal of Botany* **78**: 1164–1168.

Wardle P. 1978. Origin of the New Zealand mountain flora, with special reference to trans-Tasman relationships. *New Zealand Journal of Botany* **16**: 535–550.

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

Waser NM, Price MV. 1985. The effect of nectar guides on pollinator preference – experimental studies with a montane herb. *Oecologia (Berlin)* **67**: 121–126.

Waser NM, Chittka L, Price MV, Williams N, Ollerton J. 1996. Generalization in pollination systems and why it matters. *Ecology* **77**: 279–296.

Wilson P. 1995. Selection for pollination success and the mechanical fit of *Impatiens* flowers around bumblebee bodies. *Biological Journal of the Linnean Society* **55**: 355–383.

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

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