

DESCRIBING FLOWERING SCHEDULE SHAPE THROUGH MULTIVARIATE ORDINATION

Emily J. Austen,^{1,*} Donald A. Jackson,^{*†} and Arthur E. Weis^{*}

^{*}Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S 3B2, Canada; and
[†]School of the Environment, University of Toronto, 33 Willcocks Street, Toronto, Ontario M5S 3E8, Canada

Editor: Lynda F. Delph

Premise of research. Variation in the overall shape (e.g., skew, modality) of the flowering schedule can affect intra- and interspecific interactions, but full appreciation of these effects is hindered by the difficulty of describing and comparing schedule shapes. We propose a novel approach to schedule description and comparison based on multivariate ordination. Unlike other methods of describing shape, ordination does not require a priori assumptions about the nature of schedule variation and so may detect variation otherwise unseen.

Methodology. Using illustrative data to develop the approach, we investigated chord distances and Kolmogorov-Smirnov distances as measures of pairwise differences in schedule shape in principal coordinates analysis. We applied this same technique to an empirical data set and used constrained principal coordinate analysis to determine whether variation in schedule shape in the empirical data set can be explained by other phenological variables.

Pivotal results. Principal coordinates analysis identified biologically meaningful variation in schedule shape in both the illustrative and empirical data sets. For both distance measures and both data sets, the first ordination axis arranged individuals by the skew of their flowering schedule, and the second axis separated bimodal schedules from unimodal schedules, particularly at intermediate regions on axis 1. Constrained ordination of the empirical data set revealed that schedule skew tended to vary with flowering duration and timing of flowering onset and that plants producing more flowers tend to be more unimodal in their flowering schedule.

Conclusion. Multivariate ordination successfully separated individuals according to the shape of their flowering schedule and identified relationships between phenological variables of potential biological importance. The relationship between onset and skew, for example, could affect the realized strength of phenological assortative mating. Ordination is a departure from function-fitting schedule description. It is broadly applicable to other types of phenological data sets, offering, for example, a new tool for tracking phenological changes over years.

Keywords: assortative mating, canonical analysis of principal coordinates, flowering schedule, phenology, principal coordinates analysis.

Online enhancements: appendix figures, supplementary data file.

Introduction

The reproductive schedule is a suite of life-history traits of profound ecological and evolutionary importance. In plants, flowering time mediates an individual's interactions with mutualists (e.g., pollinators, seed dispersers), antagonists (e.g., florivores, frugivorous predators), and the abiotic environment (e.g., frost, drought; Augspurger 1981; Campbell 1991; Mahoro 2003; Griffith and Watson 2005; Inouye 2008). A full characterization of these interactions requires a quantitative

description of phenological variation (Fox 2003). Some aspects of phenology, including flowering onset, flowering end, duration of flowering, and flowering peak, can be readily measured and compared across populations (Inouye 2008) and/or across individuals within populations (Weis and Kossler 2004). Others, most notably the overall temporal pattern or "shape" of flower deployment, are more difficult to quantify and compare.

Phenological shape is of particular interest because shape can alter the probability of mating between individuals. Two individuals cannot mate if their flowering schedules do not overlap. Consequently, early-flowering individuals tend to mate disproportionately with other early individuals and late with late. This phenological assortative mating inflates genetic variance of the assortative trait and correlated traits (Weis et

¹ Author for correspondence; e-mail: emily.austen@mail.utoronto.ca.

Manuscript received June 2013; revised manuscript received September 2013; electronically published December 5, 2013.

al. 2005) and can thereby modify response to selection (Fox 2003). Schedule shape, in turn, can temper the strength of phenological assortative mating. If, for example, two individuals fully overlap in their timing of flowering but differ in the skew of their flowering schedule, mating opportunity between them will be less than predicted by perfectly matched schedules (Fox 2003). This cryptic component of phenological assortative mating can be detected only by comparing phenological shapes.

The shape of the flowering schedule can also affect interactions at the community level. Patterns of flower deployment can influence resource availability to pollinators. When community flower production is bimodal, such that most flowers are produced early or late in the season, pollinators may suffer midseason nectar and pollen scarcity (Aldridge et al. 2011). The prevalence and severity of bimodal flowering distributions may increase with climate warming in some communities (Sherry et al. 2007). To monitor these types of community changes, a measure of phenological shape is required.

Available tools for describing and comparing phenological shapes are all somewhat limited. Metrics such as the day of median flowering (Buide et al. 2002) or the distance between median flowering and the midpoint of an individual's flowering duration (Weis and Kossler 2004) are problematic because they can confound shapes. In figure 1, for example, individuals A and J would be indistinguishable by either of these metrics, despite varying markedly in schedule shape. Modeling shape by a function, such as exponential sine (Malo 2002), Weibull (Brown and Mayer 1988), logistic growth (Meagher and Delph 2001), or epsilon-skew-normal curve (Clark and Thompson 2011), is similarly problematic because functions require a priori assumptions; e.g., phenology modeled by any of these functions is restricted to a unimodal distribution. Curve fitting using spline functions (Moussus et al. 2010) offers flexibility in the shape being fitted, but two phenological distributions must be of the same overall shape for meaningful comparisons between schedules to be made using this method.

We propose a novel approach to quantifying flowering schedule shape based on multivariate ordination. Multivariate ordination simplifies complex data sets (e.g., those recording flowers per plant over a number of days) by extracting new variables (axes) that describe principal patterns of variation in the data set and by ordering observations in the reduced space described by these new variables (Legendre and Legendre 1998). Similarities or differences between pairs of observations (i.e., pairs of plants) in the original data set are summarized in the reduced ordination space. Principal components analysis (PCA), a multivariate ordination technique familiar to many ecologists, preserves pairwise Euclidean distances in the original data. We use a related technique, principal coordinates analysis (PCoA), that accommodates a range of metric (and in some cases semimetric) distance measures. Our goals are (1) to develop the multivariate ordination approach in a worked example, (2) to apply ordination to an empirical data set, and (3) to evaluate the effects of predictor variables on schedule shape in the empirical data set. To our knowledge, this is the first application of multivariate ordination to the description of phenological variation. We aim to lay out a methodological framework for additional development.

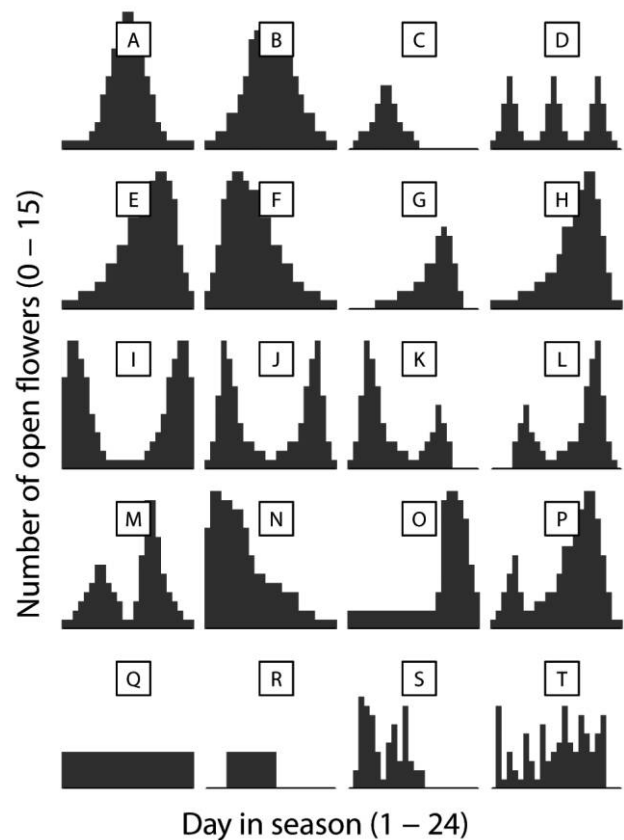


Fig. 1 Bar plots depicting schedules of flower deployment for 20 simulated flowering schedules in the illustrative data set. These schedules were chosen to capture a range of biologically interesting variation in the shape of the schedule of flower deployment.

Material and Methods

Overview of Analytical Approach

We begin with an $n \times p$ flowering data set X , with one row for each of n plants on which flower observations have been made and one column for each of p days of flower observations. This is typical of phenological data collected in field and greenhouse studies (Brunet 1996; Weis and Kossler 2004; Haggerty and Galloway 2010; Ishii and Harder 2012). Cell X_{ij} is the number of open flowers on individual i on day j of the season; if i presented no flowers on day j , $X_{ij} = 0$. From this data set, a new $n \times n$ distance matrix is developed, reporting the difference (“distance”) in flowering schedules between each pair (a, b) of individuals.

We consider the calculation of distance between schedule shapes a and b below, but in general, PCoA works well on metric distances, i.e., those meeting the following criteria (Legendre and Legendre 1998): (1) if $a = b$, then distance $ab = 0$, (2) if $a \neq b$, then distance $ab > 0$, (3) distance $ab =$ distance ba , and (4) distance $ab +$ distance $bc >$ distance ac (triangle inequality). A useful PCoA solution can sometimes be obtained when condition 4 is violated.

In a manner similar to PCA, PCoA summarizes patterns of

variation in the distance matrix by extracting a series of uncorrelated axes (eigenvectors). Subsequent axes represent successively reduced amounts of variation, with variation across all axes summing to 100% of the variation contained in the distance matrix when a metric distance measure is used. Up to $n - 1$ axes can be extracted (see Legendre and Legendre 1998; Hirst and Jackson 2007 for details and examples). Because the proportion of variation explained decreases from the first to last axis, generally only the first few axes will be informative. The goal is to summarize variation associated with many variables into these first few uncorrelated axes for graphical interpretation or use in subsequent analyses.

When the original observations are plotted on the new ordination axes, an individual's score along the ordination axes quantitatively describes its flowering schedule shape. Pairs of observations with a large pairwise distance in the $n \times n$ distance matrix will be positioned at a large distance from one another in the reduced ordination space, and those with a small pairwise distance will appear close together. If a metric distance measure has been used, pairwise Euclidean distances along all PCoA axes will equal pairwise distances in the distance matrix.

The above outlines a purely descriptive, unconstrained PCoA; the axes extracted are those that maximize the variation in schedule shape among observations. A related, constrained PCoA technique, termed canonical analysis of principal coordinates (CAP; Anderson and Willis 2003), can be used to ask whether schedule shape is affected by a given set of predictor variables. CAP is analogous to constrained PCA techniques such as canonical correlation analysis or discriminant analysis (Anderson and Willis 2003). CAP begins with the same distance matrix used in PCoA, but its solution is constrained to maximize the correlation between the extracted axes and predictor variables. CAP returns a series of "constrained axes" equal to the number of predictors in the model and a series of "unconstrained axes" accounting for variation not summarized on the constrained axes. The constrained axes may or may not correspond to the first axes identified through PCoA, the outcome depending on whether the chosen predictor variables correlate to the major axes of variation in the data set (Anderson and Willis 2003). The unconstrained PCoA and the constrained CAP analyses offer complementary views of patterns in multivariate space, the former describing multivariate variation and the latter relating that variation to hypotheses. We ran PCoA and CAP in R (R Development Core Team 2012), using package *vegan* (Oksanen et al. 2012), and provide scripts and sample data for analysis in a supplementary data file, available online.

Because we are seeking to describe differences in schedule shape and not differences in traits such as onset or duration (which can easily be quantified without multivariate ordination), we intentionally exclude for each individual observations made before its first and after its last nonzero flower count in calculating pairwise distances. This introduces a challenge: the number of nonzero flower counts, i.e., the number of variables describing schedule shape, varies from individual to individual. Multivariate analysis of schedule shape therefore requires either a correction to the number of variables per individual or a distance measure that can accommodate such differences.

Illustrative Data

To characterize the ability of different distance measures to summarize biologically important aspects of schedule shape, we created an $n \times p$ data set of $n = 20$ individuals and $p = 24$ days in the flowering season (fig. 1). These 20 flowering schedules varied in modality (fig. 1B, 1J, 1D), skew (fig. 1E, 1F), autocorrelation (fig. 1A, 1T), and duration (fig. 1B, 1C, 1E, 1G). From these data, we calculated chord and Kolmogorov-Smirnov (KS) distance matrices.

Chord distance. Chord distance is a metric distance measure that, in contrast to Euclidean distance, is not influenced by differences between individuals in the total number of flowers produced (Hirst and Jackson 2007). In spirit, chord distance compares the proportion of total flowers deployed by one individual versus another at each time point, and it can be thought of as Euclidean distance calculated on standardized axes. More technical descriptions are provided by Orloci (1967) and Legendre and Legendre (1998). Calculating chord

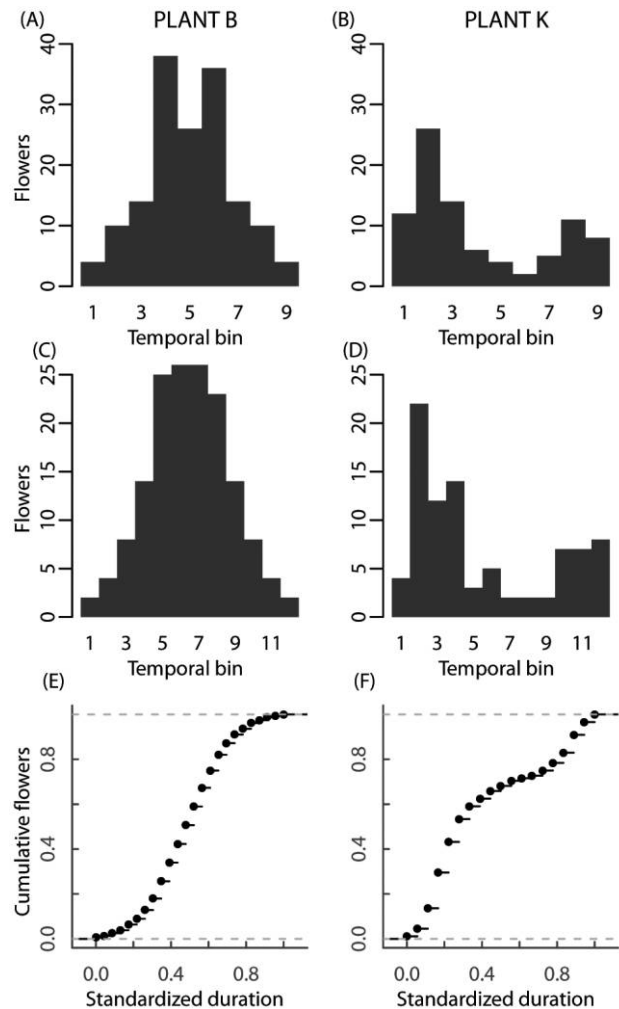


Fig. 2 Depiction of two of the 20 simulated flowering schedules (plants B and K) when standardized for calculation of chord distances based on 9 (A, B) and 12 (C, D) temporal bins and for calculation of Kolmogorov-Smirnov distances (E, F).

Table 1
Summary of Three Principal Components Analysis (PCoA) Solutions on Illustrative Data

	Chord 9		Chord 12		Kolmogorov-Smirnov ^a	
	Eigenvalue	Percent explained	Eigenvalue	Percent explained	Eigenvalue	Percent explained
PC1	1.99	34.6	2.10	34.2	.60	49.4
PC2	1.88	32.8	1.95	31.8	.31	25.5
PC3	.76	13.3	.61	9.9	.10	8.5
PC4	.49	8.5	.53	8.7	.07	5.5
Deviation (%) ^b	0		0		10.1	

Note. Chord 9 and chord 12 refer to solutions calculated on chord distance matrices based on dividing flower schedules into 9 and 12 temporal bins, respectively.

^a Percent variation explained for PCoA solution on Kolmogorov-Smirnov distances are adjusted to include only positive eigenvalues when calculating the total variation explained.

^b Deviation (%) is calculated as the ratio of the sum of negative eigenvalues to the sum of positive eigenvalues, multiplied by 100.

distances required a correction to differences in the number of nonzero days describing individuals. We divided each individual's nonzero flowering duration into q equally sized temporal bins and summed the number of flowers occurring within bins, replacing p days of flowering observations with q temporal units (supplementary data file). The number of days per bin varied from individual to individual depending on flowering duration, but q , the number of bins per individual, was constant across all plants (fig. 2A–2D). Conceptually, individuals are plotted in q -dimensional space, with position along a given axis equal to the proportion of flowers presented during that time interval. Chord distance is a function of the angle between vectors drawn from the origin to two individuals in this q -dimensional space. Where two observations lie on the same vector, chord distance = 0, and the maximum possible chord distance is $q^{1/2}$ (Legendre and Legendre 1998). We calculated chord distances on $q = 9$ and $q = 12$ temporal units. We chose these values by drawing histograms on the 20 simulated individuals and identifying the mode (9) and maximum (12) number of bins selected by the algorithm of function `hist()` in R.

Kolmogorov-Smirnov distance. We applied the two-sample KS measure to compare flowering schedules to one another (Sokal and Rohlf 1995). The KS statistic is the maximum distance between the cumulative flower production distributions of two individuals and is sensitive to differences in a wide range of parameters, including skew, central tendency, and dispersion (Sokal and Rohlf 1995). It ranges from 0 to 1. Because the R implementation of the KS statistic is also sensitive to differences in location, we range standardized observations so that each individual's flowering period ran from 0 to 1 (fig. 2E, 2F). Unlike chord distance, KS can accommodate differences in the number of variables (days) describing individuals.

KS distances do not meet the triangle inequality condition of metric distances. Consequently, not all variation in the distance matrix can be summarized by ordination. A PCoA solution can nonetheless be obtained, but the first PCoA axes will portray only the Euclidean portion of the distance matrix (Legendre and Anderson 1999). Remaining variation that cannot be converted to Euclidean distance is carried on later, imaginary, negative PCoA axes. The relative magnitude of positive

versus negative axes can be used to assess the reliability of the solution (Podani and Miklos 2002). Alternatively, the solution can be adjusted to remove negative eigenvalues (Legendre and Legendre 1998).

Interpretation of PCoA solutions. We generated three PCoA solutions: two based on chord distances ($q = 9$ and $q = 12$) and one on KS distances. We scaled all ordination axes to the square root of their variance explained. We assessed the usefulness of two-dimensional solutions by superimposing a minimum spanning tree calculated from the original distance matrices onto the ordination plots resulting from each solution. Crossing branches in the tree, or a lack of connection between individuals that appear close together, would suggest that the two-dimensional representation is not fully summarizing differences among individuals. We also examined the correlation between pairwise distances in the $n \times n$ distance matrix and pairwise Euclidean distances along the first two ordination axes (Legendre and Legendre 1998). Finally, for the nonmetric KS distance measure, we assessed the Euclidean representation of the distance matrix by examining the ratio of the sum of negative eigenvalues to the sum of positive eigenvalues (Podani and Miklos 2002). Large ratios suggest that a large portion of the variation in the original distance matrix could not be represented in Euclidean space.

Application to Experimental Data

Our illustrative data set is composed of “ideal” flowering schedules with little random variation and is designed to capture a wide range of shape variation. Empirical flowering schedules are probably noisier, and individuals in a single population or experiment may not vary so much in schedule. We therefore tested the utility of ordination for an empirical data set.

Experimental setup. In summer 2009, we reared 206 *Brassica rapa* (Brassicaceae) plants in a glasshouse at the University of Toronto, Canada. *Brassica rapa* is a self-incompatible annual producing an indeterminate, branched, racemose inflorescence (Gulden et al. 2008). Seeds were planted in 164-mL “conetainer” pots (Stuewe and Sons, Corvallis, OR) filled with a 3 : 1 mixture of potting mix to sand. In accordance with the goals of a larger study, we manipulated fertilizer avail-

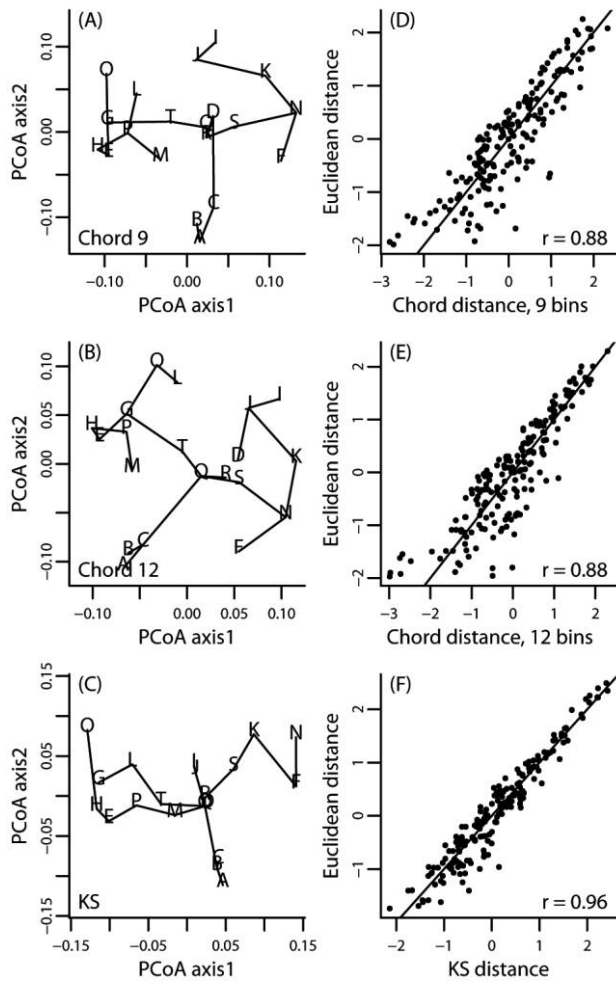


Fig. 3 Unconstrained ordination solutions with superimposed minimum spanning trees for illustrative data. Solutions are shown for chord distances where individuals are simplified into 9-bin flowering schedules (A) and 12-bin flowering schedules (B) and for Kolmogorov-Smirnov (KS) distances (C). Letters in A–C refer to individuals in the illustrative data set; ordination axes in these panels are scaled to their proportion of variation explained. Minimum spanning trees are calculated on the original distance matrix. D–F, Correlation between pairwise distances in the distance matrix and pairwise distances on the first two axes of each of the ordination solutions. PCoA = principal coordinates analysis.

ability (levels “high” and “low”) and pollination (levels “no pollination” and “full pollination”) in a 2×2 factorial design. These manipulations add interest to this analysis because resource availability and pollination may affect flowering schedules (Meagher and Delph 2001). Fertilizer treatment did not affect the timing of flower bud production (bolting time; two-sample *t*-test, $t = 0.34$, $df = 203$, $P > 0.7$).

Data collection. We recorded bolting day and the first and last day of flowering of each plant and calculated duration by subtracting the first flowering day from the last. Over the 82-d flowering period, we counted fresh, nonwilted flowers per plant every second day. Flowering schedules are thus described

in an $n \times p$ data set, where $n = 206$ individuals and $p = 41$ days on which flowers were counted. Row sums estimate total flowers produced.

Analysis. As for our illustrative data set, we range standardized each individual’s flowering schedule, calculated pairwise distances, and derived major axes of schedule shape variation through PCoA. We used CAP analysis to test the hypotheses that (1) fertilizer and/or pollination treatments and (2) other phenological characteristics, such as flowering onset and duration, affect schedule shape. CAP solutions were derived from all axes with positive eigenvalues, weighted by eigenvalue.

Results

Illustrative Data

Substantial variation was explained on the first two axes of all three PCoA solutions (table 1). Because of the violation of the triangle inequality criterion, the PCoA on KS distances returned some axes with negative eigenvalues, although their magnitude was not large relative to that of positive axes (table 1). We adjusted the denominator in our calculation of proportion of variation explained per axis to include only those axes with positive eigenvalues (table 1).

Analyses of both chord and KS distances separated individuals reasonably well, and range standardization successfully isolated variation in schedule shape from variation in day of flowering onset and duration (fig. 3A–3C). All three analyses, for example, showed that individuals A and C share a similar schedule shape despite differences in duration (fig. 3A–3C). Both KS and chord analyses characterized differences in skew, placing individuals E and F at opposite ends of the ordination space. Moreover, both succeeded in distinguishing among unimodal, bimodal, and trimodal distributions (see relative placement of individuals A, I, and D in fig. 3A–3C). Individuals exhibiting both skew and modality (e.g., K, L) are displaced on both axis 1 and axis 2 relative to nonskewed unimodal (plant A) or trimodal (plant D) schedules (fig. 3A–3C). In general, the first axis of each solution separated individuals by the skew of their flowering schedule and the second by schedule modality, particularly at intermediate values on axis 1. Overall, the three PCoA solutions were fairly congruent in their relative positioning of individuals (fig. 3A–3C).

Minimum spanning trees suggest that the two-dimensional ordination plots represent the distance matrices well (fig. 3A–3C). There is minimal crossing of branches in the solutions on chord distances and none in the solution on KS distances. Moreover, the correlation between original pairwise distances and pairwise Euclidean distances in the reduced-ordination space was strong for all three solutions (fig. 3D–3F).

Empirical Data

Because KS solutions were similar to the chord distance solutions in our illustrative data set but do not require simplification of the original data into a fixed number of bins, we used a KS distance matrix to describe our empirical data set.

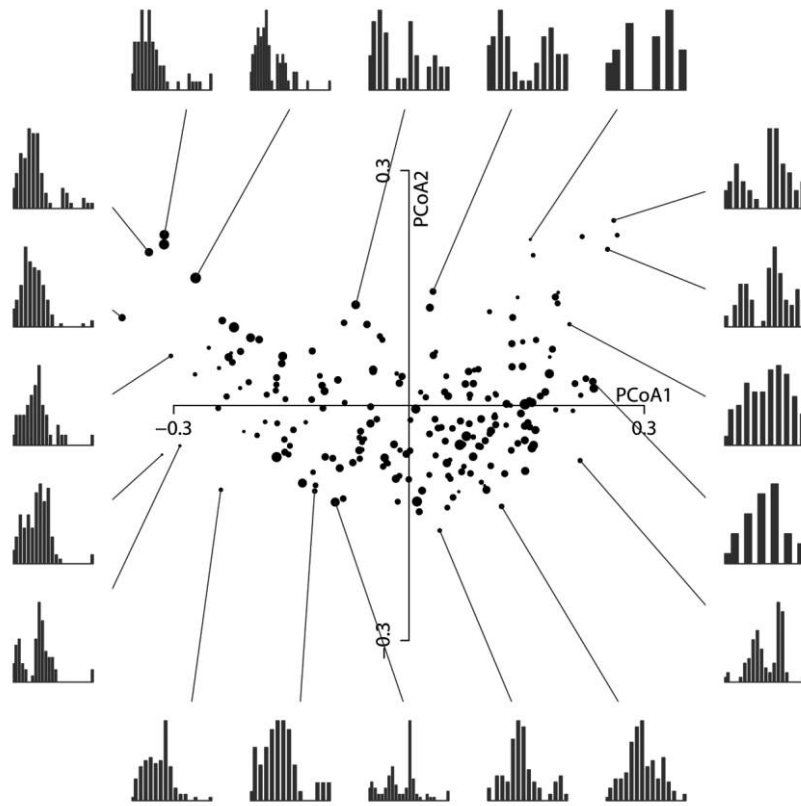


Fig. 4 Unconstrained ordination of Kolmogorov-Smirnov distances from an empirical data set of 206 plants. Point size is in proportion to the relative position of observation on axis 3. Axes in the ordination plot are scaled to their proportion of variation explained. Bar plots surrounding the central panel depict the observed flowering schedule of the observations joined to the bar plots by a line. PCoA = principal coordinates analysis.

The first PCoA axis, explaining 32.8% of the variation in the data set (adjusted to include only positive eigenvalues in the denominator) again corresponded to skew (fig. 4). Correlation between an individual's position on this axis and its skew measured as the third moment of the distribution was strong ($r = 0.92$, $n = 206$; skew calculated using R package moments; Komsta and Novomestky 2012). The second PCoA axis explained 10.0% of schedule shape variation and, by visual inspection, corresponded roughly to modality, particularly at intermediate values on axis 1 (fig. 4). Flowering schedules with positive values on axis 1 tended to be more bimodal, while those with more negative values tended toward unimodality. Variation explained on axis 3 fell to 4.2%.

A superimposed minimum spanning tree on the ordination solution (fig. A1; figs. A1–A3 available online) revealed some crossing of branches, suggesting that the two-dimensional solution was not fully capturing the variation in the original distance matrix. However, the crossing did not much affect the ordering of individuals on axis 1 or axis 2. Moreover, the correlation between KS distances and Euclidean distances on the first two PCoA axes was strong ($r = 0.92$, $n = 21,321$ unique pairs of observations), showing only minor distortion at low pairwise distances (fig. A2). The variation associated with negative eigenvalues was 39.0%.

Neither fertilizer treatment nor pollinator treatment emerged as important predictors of schedule shape variation in CAP analysis (table 2). We found stronger relationships between schedule shape and bolting date, duration, and total number of flowers produced (table 3; fig. 5). To confirm that the proportion of variation explained by these other phenological variables exceeded that expected by chance, we permuted the row (plant) order in the matrix of predictor variables, recalculating the CAP solution for each permutation. The variation explained by constrained axes when the data set is in the correct order (15.8%; table 3) far exceeded that explained by chance (mean \pm SE = 1.47% \pm 0.01%; one-tailed t -test, $t = -1106$, $P < 0.01$, $n = 1000$ permutations).

The first axis of this constrained ordination correlated positively with the first axis of unconstrained PCoA ($r = 0.99$) and the second with the second ($r = 0.93$). Later bolting and longer flowering duration were both associated with more negative values on axis 1 (i.e., more negative skew; fig. 5B). Longer flowering duration was also associated with more negative values on axis 2 (fig. 5B). Total flower production loaded negatively on axis 2 (fig. 5B), meaning bimodality tended to be more pronounced in plants making fewer flowers. Later bolting related positively to axis 3 of the constrained solution (table 3), but this axis contributed little to explained variation and

Table 2

Summary of Axes Generated through Constrained Ordination (CAP) of Kolmogorov-Smirnov Distance Matrix of Empirical Data Using Fertilizer Treatment and Pollination Treatment as Predictors of Variation in Schedule Shape

	CAP1 ^a	CAP2 ^a
Constrained axes:		
Fertilizer (high)	-.69	-.72
Pollination (full)	.71	-.70
Eigenvalue	.19	.06
Percent explained	1.4	.5

Note. Total inertia = 8.01; adjusted inertia (corrected for imaginary axes) = 13.13 (100%); inertia constrained axes = 0.25 (1.9%); inertia unconstrained axes = 12.89 (98.1%); inertia imaginary axes = -5.12. CAP = canonical analysis of principal coordinates.

^a Reported loadings are the raw output and have not been scaled to the proportion of variation explained by the axis.

we were not able to discern the nature of schedule shape variation along it.

Discussion

Through analysis of illustrative and empirical data, we have demonstrated that variation in the overall shape of the flowering schedule can be quantitatively summarized and interpreted through PCoA and can be related to variation in predictor variables through CAP. Below, we discuss the advantages and limitations of the ordination approach and offer an interpretation of the placement of observations in the KS PCoA space. We discuss the CAP results in the context of standard and cryptic assortative mating and correlated selection on phenological traits. We conclude by suggesting other possible applications of ordination of schedule shape.

Advantages and Limitations

Schedule description through multivariate ordination can be preferable to description through metrics, functions, or spline fitting. This is not to say that ordination should replace these other techniques but, rather, that it adds to the suite of tools available. Unlike other approaches, PCoA allows the use of various resemblance measures that do not make a priori assumptions about the nature of shape variation and can therefore identify variation that might otherwise go undetected. Metrics and functions, by comparison, will be useful in examining one specific component of shape, perhaps after examining total schedule variation through ordination. Similarly, CAP provides an appropriate analytical approach to address questions on the effects of predictors on overall shape without a priori definition of how shape should be described.

Of course, any method can describe variation in shape only where such variation actually exists. This seemingly trivial point serves as a reminder that a plant must have produced enough flowers over enough days for a shape of flower deployment to emerge, and some species will therefore not be

amenable to schedule description through ordination or any other technique.

Other limitations are particular to multivariate ordination. First, PCoA describes variation within a given data set. Quantitative values for schedule shape obtained in the analysis of one data set cannot, therefore, be directly compared to those obtained in another. This limitation can be overcome by bringing the raw data together into a single data set for PCoA and/or CAP. The investigator may first examine whether the ordination plot is affected by differences in sampling effort, and so on, between the data sets.

Second, and more important, is the challenge (and associated advantage) of choosing the distance measure. While we found high congruency between PCoA solutions based on chord and KS distance matrices for our illustrative data (fig. 3A–3C), some discrepancy between the solutions was nonetheless apparent (fig. 3A–3C). Thus, the ordination solution (and, therefore, description of schedule shape) will vary to some degree with the distance measure used.

We adopted the solution on KS distances because it does not require simplifying schedules into a fixed number of bins. KS distance, however, is semimetric, and so not all of the variation captured by this measure can be expressed in Euclidean ordination space. While negative eigenvalues associated with semimetric distance measures can present difficulties to hypothesis testing (Legendre and Anderson 1999), they do not interfere with CAP analysis (Anderson and Willis 2003). A semimetric distance measure can be made metric by addition of a constant to the distance matrix (Legendre and Anderson 1999), but we found that such adjustment effectively resulted in the same ordination of plants in the PCoA space and did not improve our understanding of the nature of shape variation. We therefore worked with the unadjusted KS matrix. Of course, a number of other distance measures are available, and others could be developed. We encourage exploration of measures beyond those we have presented here.

Interpretation of PCoA Space for KS Distances

In both our illustrative and empirical data sets, KS PCoA axis 1 corresponded to schedule skew, and axis 2 corresponded

Table 3

Summary of Axes Generated through Constrained Ordination (CAP) of Kolmogorov-Smirnov Distance Matrix of Empirical Data Using Day of Bolting, Flowering Duration, and Log-Transformed Number of Flowers Produced as Predictors of Flowering Schedule Variation

	CAP1 ^a	CAP2 ^a	CAP3 ^a
Constrained axes:			
Bolting day	-.347	+.252	.903
Flowering duration	-.830	-.502	-.244
No. flowers (log)	-.001	-.997	-.083
Eigenvalue	1.74	.28	.06
Percent explained	13.2	2.1	.4

Note. Total inertia = 8.01; adjusted inertia (corrected for imaginary axes) = 13.13 (100%); inertia constrained axes = 2.07 (15.8%); inertia unconstrained axes = 11.05 (84.2%); inertia imaginary axes = -5.12.

^a Reported loadings are the raw output and have not been scaled to the proportion of variation explained by the axis.

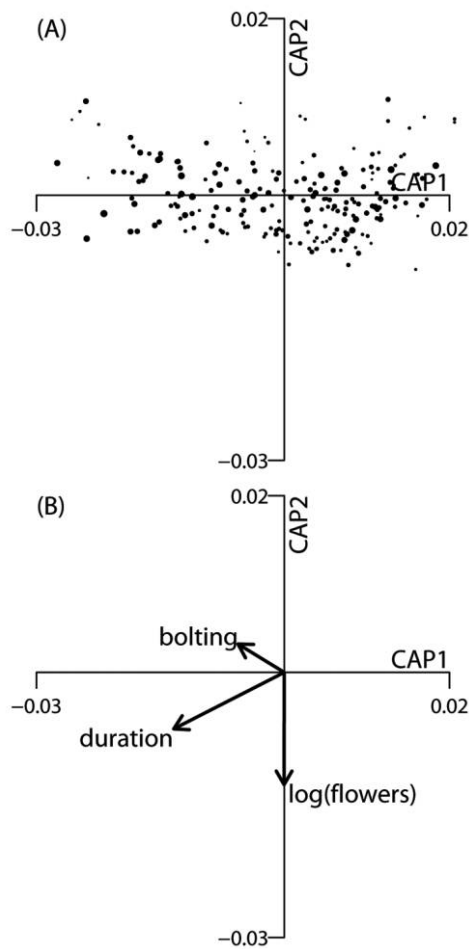


Fig. 5 Constrained ordination of Kolmogorov-Smirnov distance matrix from an empirical data set of 206 plants. *A*, Distribution of observations in constrained space; point size reflects relative positioning of observations on axis 3. *B*, Loading of predictor variables—day of bolting, total number of flowers produced (log transformed), and flowering duration—onto the first two constrained ordination axes. Axes are scaled to the proportion of variation explained. CAP = canonical analysis of principal coordinates.

roughly to modality, especially at central values on axis 1 (figs. 3C, 4). Here, we offer a more nuanced interpretation. A slight arching of the distribution of observations is evident in the empirical ordination plot (fig. 4) and, to a lesser extent, in the ordination of illustrative data (fig. 3C). Such arching is often interpreted as evidence of distortion in the two-dimensional solution due to an underlying background environmental gradient (Podani and Miklos 2002). The partial alleviation of arching in the CAP solution (fig. 5A) suggests that “environmental variables”—in this case, the environment is the plant on which the flowers have been deployed, and the variables are other phenological parameters of this plant—may indeed be contributing to the arching. However, the low proportion of variation explained on axis 3, combined with a largely correct ordering of individuals on axis 1 (figs. 3C, A1) and strong correlation between pairwise KS distances and Euclidean dis-

tances on the first two PCoA axes (fig. A2), suggests any distortion is minor. In general, as long as such arched patterns are recognized, the appropriate interpretation can be achieved.

We suggest that the graphical outcome in the PCoA ordination also reflects the reality that some regions of the PCoA space cannot be occupied because some schedule phenotypes are implausible. A flowering schedule cannot, for example, be both highly skewed and highly bimodal because the occurrence of a valley in the flowering schedule followed by a second peak

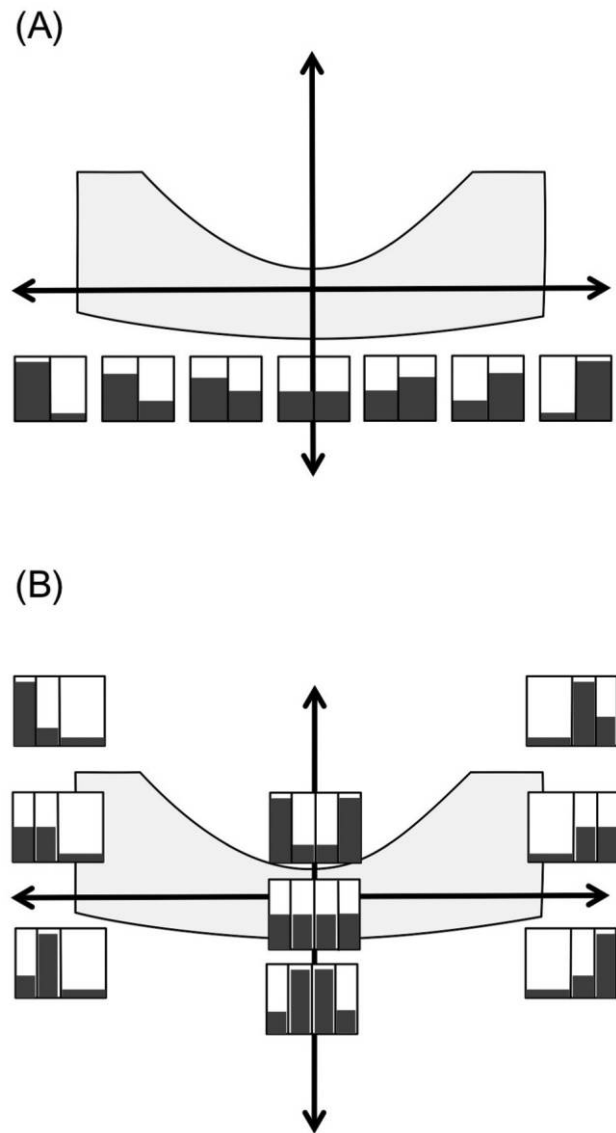


Fig. 6 Interpretation of the unconstrained ordination space defined by Kolmogorov-Smirnov (KS) distances. *A*, Because flowering schedule skew is the largest source of pairwise distances, principle coordinates axis 1 separates individuals by skew. *B*, Residual KS distances fall on axis 2; at neutral regions on axis 1, the major source of residual KS distance is whether the majority of observations occur at the center or ends of a distribution, i.e., unimodality versus bimodality. At extreme values on axis 1, residual KS distances are attributable to variation in skew not captured on axis 1.

would reduce skew. Bimodality therefore “pushes” an observation away from extreme values on PCoA axis 1 and back toward the center of the plot. See, for example, the relative positioning of individuals L and H in figure 3C. Differences in skew generate large KS distances (fig. A3). Once this major source of variation has been accounted for on axis 1 (fig. 6A), the next major sources of variation (summarized on axis 2) will be modality and/or any remaining variation in skew (fig. A3). As noted above, bimodality is most pronounced in central regions on axis 1 (fig. 6B), with bimodality detected by the density of flowers at the tails versus in the middle of an individual’s flowering period. At the extremes of axis 1, strong skew precludes bimodality, and pairwise distances (and, thus, position on axis 2) are attributed to remaining differences in skew (fig. 6B). Pairwise KS distance owing to slight differences in skew can, in fact, be comparable in magnitude to those owing to differences in modality (fig. A3). We suggest that at extreme values on axis 1, large negative values on axis 2 arise when the shape of the peak of the flowering distribution is slightly left skewed, and large positive values arise when the shape of the peak is slightly right skewed (fig. 6B).

From our interpretation of the PCoA space (fig. 6), the pairwise distances between plants at the upper right region and lower left region of the PCoA space should be smaller than the Euclidean distance in this two-dimensional ordination plot would imply. We would therefore predict that these regions should be brought closer together on axis 3. This prediction is confirmed in figure 4, where plotting of axis 3 via point size reveals that these two regions are both characterized by small values on axis 3.

The relatively low proportion of variation explained by axis 2 in the KS solution should not be taken as an indication that axis 2 is biologically unimportant. Rather, it indicates that variation in skew generates the largest pairwise KS distance (fig. A3) and that this measure is very sensitive to differences in skew. Our choice to scale the plotting axes by their proportion of variation explained (figs. 3C, 4, 5A, 5B) aided in our interpretation of the KS space. If, however, an investigator were particularly interested in differences along axis 2 (unimodal vs. bimodal schedules; Aldridge et al. 2011), it would be acceptable to consider the pattern on each axis independent of other axes. Such scaling and interpretation would imply that all axes of shape variation are of potentially equal biological importance. The choice of scaling depends on the goals of a particular study (Peres-Neto and Jackson 2001).

Multivariate Nature of Phenology

Through CAP analysis, we concluded that counter to expectations, fertilizer and pollination treatments did not affect flowering schedule shape in *Brassica rapa* (table 2). CAP analysis further revealed that variation in schedule shape in our empirical data set was not independent of variation in bolting date, duration, or total number of flowers produced (fig. 5). In *B. rapa*, late-flowering plants produce more leaves along their primary axis and therefore more axillary meristems from which to initiate secondary inflorescences than do early-flow-

ering plants (Koornneef et al. 1991; Schranz and Osborn 2000). It could be that these architectural differences allow a longer tail of flower deployment in later-flowering plants: early-flowering plants simply do not have the same opportunity for sustained flower production. Constraints arising from these types of life-history trade-offs are unlikely to evolve (Conner et al. 2011). Thus, if this hypothesis were correct, we would expect similar relationships between bolting date and ordination axes describing skew in other plants with similar growth forms.

Regardless of its underlying cause, the nonindependence among phenological variables in our data set may be biologically important. Consider again the negative correlation between day of flowering onset and CAP axis 1 (fig. 5B). This correlation could reduce the strength of phenological assortative mating relative to that arising if schedules were not skewed or if skew varied randomly with day of flowering onset. Conversely, a positive relationship between flowering onset and CAP axis 1 (skew) would tend to enhance phenological assortative mating. Given that assortative mating modifies selection response (Fox 2003; Weis et al. 2005), correlations affecting its strength will be important to the study of selection in contemporary populations.

Further Applications of the Ordination Approach to Phenological Data

We have used ordination to describe variation in the shape of flower deployment schedules among plants within simulated and experimental populations. The same approach could be used to describe patterns of variation in flowering onset or overall flower production among species within communities, or among populations of a species, or among years within a population. Through CAP, schedule variation among populations could be examined in relation to environmental variables. A data set describing several species within a community could be analyzed to explore the effects of inflorescence architecture and/or phylogeny on the flower deployment schedule. Further applications are no doubt possible. We suggest that the ordination approach will complement and enhance existing methods of describing schedule shape, perhaps the most elusive component of phenological variation.

Acknowledgments

We thank J. Forrest for constructive comments early in the development of this project; S. C. H. Barrett, J. R. Stinchcombe, J. D. Thomson, F. H. Rodd, and J. L. Ison for comments on earlier versions of this work; B. Hall and A. Petrie for plant care advice; and B. Petro, K. Turner, S. Hsieh, and S. Wadgyar for assistance in data collection. This work was supported in part by a Natural Sciences and Engineering Research Council of Canada (NSERC) graduate scholarship and an Ontario Graduate Scholarship in Science and Technology to Emily J. Austen and by an NSERC Discovery Grant to Arthur E. Weis.

Literature Cited

- Aldridge G, DW Inouye, JRK Forrest, WA Barr, AJ Miller-Rushing 2011 Emergence of a mid-season period of low floral resources in a montane meadow ecosystem associated with climate change. *J Ecol* 99:905–913.
- Anderson MJ, TJ Willis 2003 Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84:511–525.
- Augsburger CK 1981 Reproductive synchrony of a tropical shrub—experimental studies on effects of pollinators and seed predators on *Hybanthus prunifolius* (Violaceae). *Ecology* 62:775–788.
- Brown RE, DG Mayer 1988 Representing cumulative germination. 2. The use of the Weibull function and other empirically derived curves. *Ann Bot* 61:127–138.
- Brunet J 1996 Male reproductive success and variation in fruit and seed set in *Aquilegia caerulea* (Ranunculaceae). *Ecology* 77:2458–2471.
- Buide M, J Díaz-Peromingo, J Guitián 2002 Flowering phenology and female reproductive success in *Silene acutifolia* Link ex Rohrb. *Plant Ecol* 163:93–103.
- Campbell D 1991 Effects of floral traits on sequential components of fitness in *Ipomopsis aggregata*. *Am Nat* 127:713–737.
- Clark RM, R Thompson 2011 Estimation and comparison of flowering curves. *Plant Ecol Divers* 4:189–200.
- Conner JK, K Karoly, C Stewart, VA Koelling, HF Sahli, FH Shaw 2011 Rapid independent trait evolution despite a strong pleiotropic genetic correlation. *Am Nat* 178:429–441.
- Fox GA 2003 Assortative mating and plant phenology: evolutionary and practical consequences. *Evol Ecol Res* 5:1–18.
- Griffith TM, MA Watson 2005 Stress avoidance in a common annual: reproductive timing is important for local adaptation and geographic distribution. *J Evol Biol* 18:1601–1612.
- Gulden R, S Warwick, A Thomas 2008 The biology of Canadian weeds. 137. *Brassica napus* L. and *B. rapa* L. *Can J Plant Sci* 88: 951–996.
- Haggerty BP, LF Galloway 2010 Response of individual components of reproductive phenology to growing season length in a monocarpic herb. *J Ecol* 99:242–253.
- Hirst CN, DA Jackson 2007 Reconstructing community relationships: the impact of sampling error, ordination approach, and gradient length. *Divers Distrib* 13:361–371.
- Inouye DW 2008 Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology* 89:353–362.
- Ishii HS, LD Harder 2012 Phenological associations of within- and among-plant variation in gender with floral morphology and integration in protandrous *Delphinium glaucum*. *J Ecol* 100:1029–1038.
- Komsta L, F Novomestky 2012 moments: moments, cumulants, skewness, kurtosis and related tests. R package, version 0.13. <http://CRAN.R-project.org/package=moments>.
- Koornneef M, CJ Hanhart, JH van der Veen 1991 A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol Gen Genet* 229:57–66.
- Legendre P, MJ Anderson 1999 Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol Monogr* 69:1–24.
- Legendre P, L Legendre 1998 Numerical ecology: developments in environmental modeling. Vol 3. 2nd ed. Elsevier Scientific, Amsterdam.
- Mahoro S 2003 Effects of flower and seed predators and pollinators on fruit production in two sequentially flowering congeners. *Plant Ecol* 166:37–48.
- Malo JE 2002 Modelling unimodal flowering phenology with exponential sine equations. *Funct Ecol* 16:413–418.
- Meagher T, L Delph 2001 Individual flower demography, floral phenology and floral display size in *Silene latifolia*. *Evol Ecol Res* 3: 845–860.
- Moussus JP, R Julliard, F Jiguet 2010 Featuring 10 phenological estimators using simulated data. *Methods Ecol Evol* 1:140–150.
- Oksanen J, FG Blanchet, R Kindt, P Legendre, PR Minchin, RB O'Hara, GL Simpson, P Solymos, MHM Stevens, H Wagner 2012 vegan: community ecology package. R package, version 2.0-3. <http://CRAN.R-project.org/package=vegan>.
- Orlaci L 1967 An agglomerative method for classification of plant communities. *J Ecol* 55:193–206.
- Peres-Neto PR, DA Jackson 2001 The importance of scaling of multivariate analyses in ecological studies. *Ecoscience* 8:522–526.
- Podani J, I Miklos 2002 Resemblance coefficients and the horseshoe effect in principal coordinates analysis. *Ecology* 83:3331–3343.
- R Development Core Team 2012 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>.
- Schranz ME, TC Osborn 2000 Novel flowering time variation in the resynthesized polyploid *Brassica napus*. *J Hered* 91:242–246.
- Sherry R, X Zhou, S Gu, J Arnone III, D Schimel, P Verburg, L Wallace, Y Luo 2007 Divergence of reproductive phenology under climate warming. *Proc Natl Acad Sci USA* 104:198–202.
- Sokal RR, FJ Rohlf 1995 Biometry. 3rd ed. WH Freeman, New York.
- Weis AE, T Kossler 2004 Genetic variation in flowering time induces phenological assortative mating: quantitative genetic methods applied to *Brassica rapa*. *Am J Bot* 91:825–836.
- Weis AE, J Winterer, C Vacher, T Kossler, C Young, G LeBuhn 2005 Phenological assortative mating in flowering plants: the nature and consequences of its frequency dependence. *Evol Ecol Res* 7: 161–181.