JOURNAL OF EXPERIMENTAL ZOOLOGY (MOL DEV EVOL) 296B:23–39 (2003)

Evolvability and Genetic Constraint in *Dalechampia* Blossoms: Genetic Correlations and Conditional Evolvability

THOMAS F. HANSEN*[1,2], W. SCOTT ARMBRUSTER[1,3], MATTHEW L. CARLSON[1,4], AND CHRISTOPHE PÉLABON[1]
[1]*Department of Biology, NTNU, 7491 Trondheim, Norway*
[2]*Department of Biological Sciences, Florida State University, Tallahassee, Florida 32306*
[3]*Institute of Arctic Biology, University of Alaska, Fairbanks, Alaska 99775*
[4]*Alaska Natural Heritage Program, Environment and Natural Resources Institute, University of Alaska, Anchorage, 99501*

ABSTRACT The short-term evolvability of a character is closely related to its level of additive genetic variation. However, a large component of the variation in any one character may be pleiotropically linked to other characters under the influence of different selective factors. Therefore, the organization of the organism into quasi-independent modules may be an important prerequisite for evolvability. In this paper we propose to study character evolvability in terms of conditional genetic variation. By estimating the amount of genetic variation in a character, y, that is independent of other characters, x, we can assess the evolvability of y when there is stabilizing selection on x. We suggest that systematic use of conditioning may help build a picture of modular organization and quasi-independent evolvability. As an illustration, we use this approach to assess the evolvability of floral characters in *Dalechampia scandens* (Euphorbiaceae). Although our study population had relatively low levels of genetic variation at the outset, we find evidence that conditioning may lead to substantial further reduction in the genetic variation available for independent adaptation. This provides additional evidence that the *D. scandens* blossom is constrained in its short-term evolvability. *J. Exp. Zool. (Mol. Dev. Evol.) 296B:23–39, 2003.*

INTRODUCTION

Evolvability may be defined as the ability or potential to respond to selective challenges (Wagner and Altenberg, '96). In evolutionary quantitative genetics this is usually related to the additive genetic variance, which determines the direct ability of a trait to respond to a novel selection pressure over one or a few generations (e.g., Lande, '79; Houle, '92). However, using the additive genetic variance to predict a response to selection may severely overestimate evolvability, as the observed variation of a given character may not be fully available for adaptation of that character due to pleiotropic constraints. For example, an unknown fraction of the new mutational variation that appears in any one character may be due to degenerative changes in genes with general biological effects in the organism, and thus carry along a set of deleterious pleiotropic effects (Kondrashov and Turelli, '92). This may generate apparently usable variation in any one character, but is unlikely to provide a basis for permanent evolutionary change. Mutations with widespread deleterious effects may be largely weeded out of standing genetic variation, but pleiotropic constraints are still expected on the population level, in the form of genetic correlations between characters that are under discordant selection. The relative stability of organisms over evolutionary time suggests that most characters spend the majority of time in the vicinity of some local fitness optimum and are therefore normally subject to stabilizing selection (Williams, '92; Rowe and Houle, '96; Hansen, '97). A shift in the environment may

Grant sponsor: The Norwegian Research Council; Grant numbers: 123846/410, 123650/410, and 128830/410 to TFH and WSA.
*Correspondence to: Thomas F. Hansen, Department of Biological Science, Conradi Building, Florida State University, Tallahassee, FL 32306. E-mail: Thomas.Hansen@bio.fsu.edu.
Received 3 June 2002; Accepted 18 November 2002
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.00014

place some characters under directional selection until they approach a new local optimum, but this is hardly likely to affect all characters simultaneously. Under this sort of mosaic selection, the evolvability of a character is largely determined by the part of its genetic variation that is unconstrained by other characters under stabilizing selection.

Evolvability is therefore intimately related to the modularity and integration of characters (Olson and Miller, '58; Berg, '60; Riedl, '77, '78; Kauffman, '93; Cheverud, '96a, 2001; Raff, '96; Wagner, '96; Wagner and Altenberg, '96; Gerhart and Kirschner, '97; Von Dassow and Munro, '99; Stern, 2000; Magwene, 2001; Hansen, 2003). Only to the extent that a character is able to vary independently of the rest of the phenotype can it be molded by selection in such a way that it can be considered an independent adaptation (Lewontin, '78). To achieve an understanding of evolvability, it is therefore crucial to study the genetic variability of a character that remains when other aspects of the organism are not allowed to vary.

Here we propose to use estimates of conditional genetic variance to produce a better picture of the amount of variation that is actually available for adaptation. We refer to the ability of a character to evolve independently of other (specified) characters as conditional evolvability. A trait may be conditioned on characters that are presumed to be under strong stabilizing selection, so as to assess the ability to evolve without perturbing selectively constrained characters. Alternatively, it may be conditioned on developmentally related traits, or earlier developmental stages, so as to assess the effect of developmental constraints on patterns of genetic variation. Conditioning may also be conducted on functionally related characters to test whether functional relationships are reflected in the genetic architecture.

We illustrate the concept of conditional evolvability with a study of the floral morphology of a population of *Dalechampia scandens* (Euphorbiaceae), a neotropical vine pollinated by resin-collecting bees (Armbruster, '84, '85 '96). In a previous study (Hansen et al., 2003), we estimated components of variation and used this to assess the unconditional evolvability of separate floral traits. In this paper, we study character integration and ask how much the evolvability may be reduced by conditioning traits on each other.

THEORY

Conditional evolvability

Under natural conditions most characters may be constrained by genetic and phenotypic correlations with other parts of the organism that are under stabilizing selection. To characterize the component of additive genetic variation that is actually available for adaptation, we consider two multivariate characters as described by their trait vectors x and y. Let y be the character under directional selection and let x represent a character under strong stabilizing selection that we suppose constrains the evolution of y. Lande's ('79; Lande and Arnold, '83) equations for the evolutionary changes in the mean character values, X and Y, can be written

$$\Delta Y = G_y\beta_y + G_{yx}\beta_x, \tag{1a}$$

$$\Delta X = G'_{yx}\beta_y + G_x\beta_x, \tag{1b}$$

where G_y and G_x are the genetic variance matrices of y and x, respectively, G_{yx} is their genetic covariance matrix, and β_y and β_x are the selection gradients on the two characters. Assuming the G-matrix remains constant, we may sum these equations over several generations

$$\Sigma\Delta Y = G_y\Sigma\beta_y + G_{yx}\Sigma\beta_x, \tag{2a}$$

$$\Sigma\Delta X = G'_{yx}\Sigma\beta_y + G_x\Sigma\beta_x, \tag{2b}$$

If we now assume that the cumulative change in X is zero, which should be approximately true if this character is under strong stabilizing selection, we find that the cumulative change in Y is

$$\Sigma\Delta Y = (G_y - G_{yx}G_x^{-1}G'_{yx})\Sigma\beta_y = G_{y|x}\Sigma\beta_y, \tag{3}$$

where $G_{y|x}$ is the genetic variance matrix of y conditional on the genetic component of x. This conditional genetic variance matrix thus describes the evolvability of one character when another is not allowed to change. Note that it would not be legitimate to assume that there is no change in X over the first generation. Even if character x is under stabilizing selection it may change due to both indirect selection and correlated response to selection on character y. However, this would then be compensated by direct selection on x in subsequent generations making it more reasonable to assume that the cumulative change is zero. Furthermore, the assumptions of strong stabilizing selection and zero change in X are only

heuristic, and not strictly necessary, as it can be shown that $G_{y|x}$ determines the rate of response in Y whenever X is in a steady state even when X is displaced from its optimum (Hansen, 2003). Of course, unbreakable constraints would result in a trade-off where both characters are kept away from their individual optima.

It is worth pointing out that $G_{y|x}$ does contain sources of variation that are shared with x. Consider the following simple model for the genetic variances of x and y, which we now for illustration assume are univariate,

$$\begin{aligned} G_y &= g_y + g_{yx}, \\ G_x &= g_x + g_{yx} \end{aligned} \tag{4}$$

Here, g_{yx} is a component of variance that is shared between the traits, while g_y and g_x are trait-specific components of variation. If we assume these three sources of variation are independent, the conditional genetic variance becomes

$$G_{y|x} = g_y + \frac{g_{yx}g_x}{G_x} \tag{5}$$

From this we see that a certain fraction of the shared variance can also be used in a response to selection without compromising the constraining character. This fraction is equal to the percent of variance in the constraining character, x, that is not shared with y. Thus, if a trait is only sharing its variation with characters that have ample independent sources of variation, its evolvability will not be hampered. This demonstrates that our measure of evolvability allows for compensatory change in the genetic basis of the constraining traits, and that phrases such as ''variation independent of other traits'' should be interpreted in this sense.

Conditioning on genotypic versus phenotypic value

Zhu ('95) suggested estimating the amount of genetic variation at a later developmental stage that was independent of an earlier developmental stage by using a phenotypic regression of late on early stage and subject the residuals to quantitative genetic analysis. This approach can be straightforwardly generalized to any pair of (multivariate) characters. This procedure estimates the genetic component of the conditional phenotypic variance, which we will denote $G_{y|Px}$. Thus, $G_{y|Px}$ is conditional on the phenotypic, not genotypic, value of the x-character. Therefore, $G_{y|Px}$ and $G_{y|x}$ are not the same, and may be considered alternative measures of conditional evolvability.

Conditioning on the phenotypic value may provide a better prediction for selection response in a single generation, as stabilizing selection will act to keep the phenotypic rather than the genotypic mean constant. However, indirect selection will then allow a genotypic response in the x-variable by creating a negative G x E correlation. As the change in the non-genetic component of x will usually not be transferred to the next generation, we expect to find X displaced from its optimum in the next generation. Thus, the extra evolvability that this allows may be offset by selection on x in the next generation. Therefore, $G_{y|x}$ may be the better predictor of evolvability over several generations.

Although this makes $G_{y|x}$ the preferred measure of evolvability from a theoretical point of view, $G_{y|Px}$ can be estimated with better accuracy and much less computational effort (see methods). It can also be used to assess the effects of conditioning on dominance, epistatic, maternal, and environmental components of variance. In this study we report both measures.

MATERIALS AND METHODS

Materials and experimental design

The plants used for estimation of quantitative genetic parameters derive from seeds collected from 84 separate individual *D. scandens* near Tulum, Mexico. Four additional populations, one from Mexico and three from Venezuela where used to study interpopulation differences. These also are derived from field-collected seeds. The treatment of plants is described in Hansen et al. (2003).

The genetic study was designed as a block diallel where 12 sets of five parental individuals grown from the field-collected seeds were combined into 12 complete 5x5 diallels with reciprocals. Two individuals were raised from each mating; thus there where four full sibs from each parental pair.

Traits and measurements

The pseudanthial *D. scandens* blossom is composed of two large involucral bracts, usually 10 male flowers, three female flowers and a large resin-secreting gland. We measured a number of traits early in the bisexual stage of the blossom. These are shown in Figure 1 and Table 1.

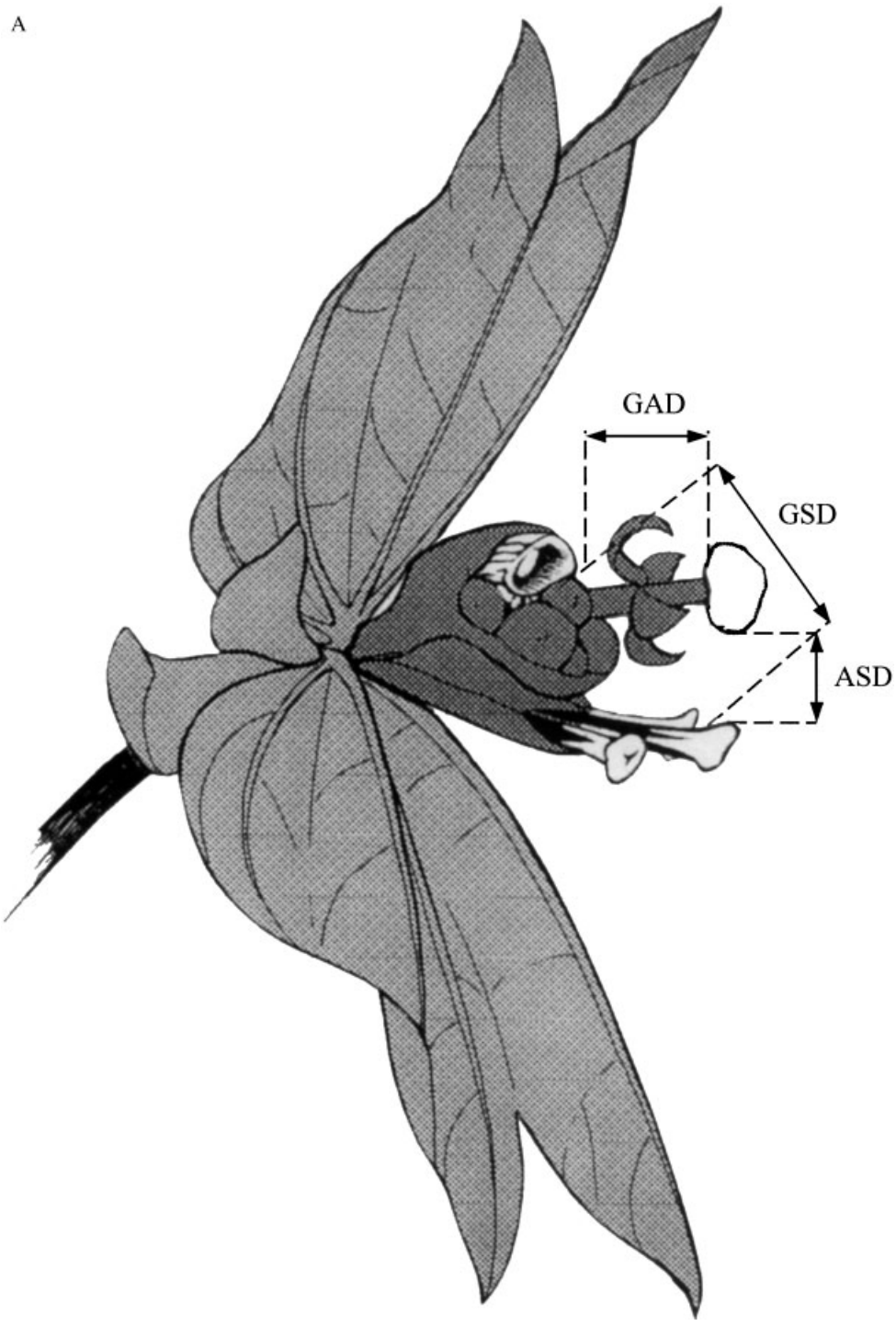

Fig. 1. A) Side-view of *Dalechampia scandens* blossom, indicating gland-anther distance (GAD), gland-stigma distance (GSD), and anther-stigma distance (ASD). B) Exploded view showing remaining floral measurements (see text for explanation).

The white-to-greenish involucral bracts open during the day to expose the flowers to pollinators and they probably have a signaling function analogous to petals. They also have a protective function, as they close over the blossom during night and are closed during development of both the bud and the maturing fruits (Armbruster, '96). We measured the width and lengths of the bracts.

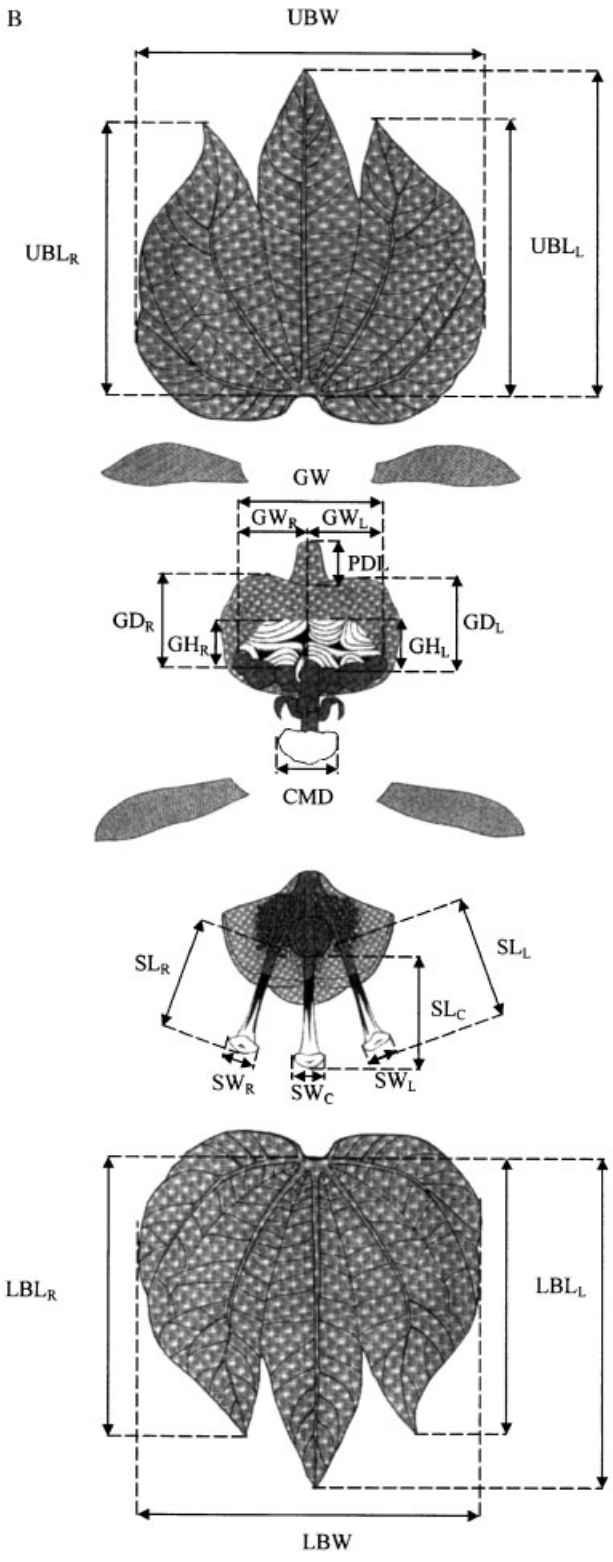

Fig. 1. *Continued*

The upper and lower bract widths are designated UBW and LBW, respectively. The upper and lower bract lengths, UBL and LBL, are averages of the lengths from the bract base to the tip of the three lobes.

The male flowers are arranged in three branches of three flowers each, surrounding a single central (terminal) flower. The central flower is always the first to open and remains the only open flower for at least one day. All measures of male flowers involve this central flower. We measured the gland-anther distance, GAD, which is important as it determines the minimum size a bee must be to receive pollen (Armbruster, '88, '90). We also measured the diameter of the stamen cluster of the central male flower, CMD, and the anther-stigma distance, ASD, the distance from the central male flower to the stigma of the central female flower. The anther-stigma distance affects the probability of self pollination (Armbruster, '88). The gland-stigma distance, GSD, is the average of the shortest distances from each of the three female stigmas to the gland and is likewise important in determining what sized bee can pollinate the blossom.

In addition to GSD, we also measured the length of the style and width of the stigmatic tip of the three female flowers. The style length, SL, and stigma width, SW, are averages of these. The style length affects GSD and ASD. The distance from gland to flowers is also affected by the size of the gland and, in case of GSD, the length of the peduncle (PDL) of the male cymule. A long peduncle will bring the gland and the male flowers closer to the stigma tip.

The gland is composed of two symmetric lobes. We measured the height, width and depth of the two lobes. Averages of the two sides were used to form the gland depth, GD, and gland height, GH. The gland width, GW, was made as a single measurement. The gland area, GA = GH $\times$ GW, is a measure of the resin-secreting surface which is correlated with the standing crop of resin that the blossom can offer to pollinators, and in turn influences what bees are attracted (Armbruster, '84). The gland is composed of some 15 to 30 resin-secreting bractlets. The number of bractlets, GN, is an important determinant of gland size.

Measures were taken by two observers. Data set 1 was measured by CP on two blossoms from each individual without dissecting the blossom. Data set 2 was measured with dissection by TFH on one blossom from each of only two individuals from each full-sib family. The traits GN, GD, CMD, SL and PDL are only available in the second data set. Except when some of these traits are involved we use data set 1 for univariate analyses and

TABLE 1. Trait definitions, means and evolvabilities: Based on analysis presented in Hansen et al. (2003)

Trait	Measurement	Mean	Evolvability (I_A)
UBW (Upper Bract Width)		20.59 ± 0.24	$0.31\% \pm 0.09\%$
UBL (Upper Bract Length)	$(UBL_L+UBL_c+UBL_R)/3$	17.55 ± 0.21	$0.25\% \pm 0.07\%$
LBW (Lower Bract Width)		20.73 ± 0.31	$0.34\% \pm 0.10\%$
LBL (Lower Bract Length)	$(LBL_L+LBL_c+LBL_R)/3$	18.61 ± 0.23	$0.28\% \pm 0.08\%$
GAD (Gland-Anther Distance)		4.66 ± 0.05	$0.12\% \pm 0.06\%$
GSD (Gland-Stigma Distance)	$(GSD_L+GSD_c+GSD_R)/3$	4.64 ± 0.06	$0.48\% \pm 0.14\%$
ASD (Anther-Stigma Distance)		3.62 ± 0.08	$1.71\% \pm 0.56\%$
CMD (Central Male flower Diam.)		2.79 ± 0.03	$0.15\% \pm 0.09\%$
GW (Gland Width)		6.61 ± 0.08	$0.11\% \pm 0.05\%$
GH (Gland Height)	$(GH_L+GH_R)/2$	2.92 ± 0.05	$0.33\% \pm 0.13\%$
GD (Gland Depth)	$(GD_L+GD_R)/2$	2.96 ± 0.03	$0.35\% \pm 0.13\%$
GN (Gland Number)	# bractlets in gland	21.22 ± 0.49	$1.46\% \pm 0.58\%$
GA (Gland Area)	GH*GW	19.56 ± 0.53	$0.78\% \pm 0.31\%$
PDL (PeDuncle Length)		3.16 ± 0.06	$0.98\% \pm 0.50\%$
SL (Style Length)	$(SL_L+SL_C+SL_R)/3$	6.27 ± 0.11	$0.49\% \pm 0.23\%$
SW (Style Width)	$(SW_L+SW_C+SW_R)/3$	1.35 ± 0.02	$0.33\% \pm 0.11\%$

Abbreviations in the measurement column refer to Figure 1. The evolvability is measured with $I_A=G_y/Mean^2(\pm SE)$.

phenotypic conditioning. We use only data set 2, however, for genetic conditioning and genetic correlations, as it was not computationally feasible to do multivariate analyses on data set 1.

Measuring evolvabilities

Ignoring constraints, the ability of a population to respond to a directional selection pressure is largely determined by the additive genetic variance. Short-term, or ''population,'' evolvability has traditionally been measured by the heritability, h^2, the ratio of additive genetic to phenotypic variance. Unfortunately, heritabilities are not suitable as measures of evolvability due to the strong correlations between additive genetic and environmental sources of variation, and the tendency for h^2 to be negatively correlated with its associated measure of selection strength, the selection differential, S. For traits on a ratio scale, a mean-standardized additive genetic variance may be a more appropriate measure of evolvability (Houle, '92). As discussed in Hansen et al. (2003), I_A, the additive genetic variance divided by the square of the trait mean, is an operational measure of population evolvability, as it, under certain assumptions, equals the percentage expected evolutionary response to a unit strength of directional selection, ϕ, where ϕ is the (invariant) strength of selection on fitness. Thus, the I_A of a trait is interpretable as the response the trait would show if selection were to be as strong as selection on fitness itself. We will use this as our measure of evolvability.

Statistics

The pedigree consisting of full sibs, half sibs, and some other types of relatives (described in Hansen et al., 2003) was analyzed statistically as described in Lynch and Walsh ('98, chaps. 26 and 27) with a mixed model implemented into PROC MIXED in SAS 6.12 by use of the TYPE = LIN general linear variance structure and restricted maximum likelihood (REML) estimation of variance components. We included an additive genetic effect in all analyses. In the analyses involving data set 1, but not data set 2, we also included stage (i.e., one, two, or three male flowers open) as a fixed effect, and, as there are two blossoms from each individual in this set, we included individual as a random effect.

Maternal and dominance effects were unimportant for all traits (Hansen et al., 2003) and were not included in the analysis. There was some temporal variation in blossom traits, but as this did not much affect the estimates of genetic variance components (Hansen et al., 2003), it was not included in the present analysis.

Results from the two data sets were largely consistent, but the estimates of additive genetic variance tended to be slightly higher in data set 2. One important trait, GA, showed much higher levels of genetic variance in data set 2. In this case,

we believe the estimate from data set 1 is the most accurate, because the estimate of selfed-sib variance from data set 2 was more consistent with the former (not shown).

Genetic covariances were estimated from data set 2 by combining pairs of traits into a single response vector and including a ''trait'' variable as both a fixed and a random effect (Searle et al., '92; Hansen and Boonstra, 2000). The interaction of trait and parents was used as random effect and three coefficient matrices were used: two corresponding to the additive genetic variances of the traits and one to their additive genetic covariance. A residual covariance was also included. Each pair of traits was analyzed separately and no other variance components were included.

Genetic variances and covariances conditional on genetic values of other traits (i.e., $G_{y|x}$-matrices) were computed from the entire G-matrix by substituting the appropriate elements into the equation $G_{y|x} = G_y - G_{yx}G_x^{-1}G_{yx}$. The genetic components of conditional phenotypic variability (i.e., the $G_{y|Px}$-matrices) were computed from data set 1 by including the x-variables as fixed effects in the mixed model, assuming the slope of the effect is the same within each stage. Including the x-variables directly as covariates in the mixed model is slightly different from Zhu's ('95) suggested analysis of residuals. Our approach allows for generalized least squares estimation of the regression slope that takes the family structure into account.

We will report the effect of conditioning as a percent reduction in variance as compared to the unconditional variance. In the case of genotypic conditioning this is equivalent to $1 - R^2$, where R^2 is the square of the multiple correlation coefficient of the genetic value of y with the genetic values of the x variables (Anderson, '84). The interpretation is similar for conditioning on the phenotype, but note that the additive genetic variance may actually increase after conditioning, as all variance components are computed anew from the residuals.

Standard errors of I_A-evolvabilities were computed using a standard approximation for the variance of a ratio (Lynch and Walsh, '98, p. 818). Standard errors of the percent reduction in evolvability due to phenotypic conditioning in Table 2, were, however, based only on the sampling error of the conditional variance, as the sampling errors of conditional and unconditional variance are likely to be strongly correlated.

As conditional genetic variances are estimated as functions of REML estimated variance components, they are themselves maximum-likelihood estimators. This because the maximum-likelihood estimate of a function of parameters is equal to the same function applied to the maximum-likelihood estimates of the parameters (Anderson, '84). This guarantees consistency, as well as asymptotic normality and efficiency of the estimator. Genetic variances conditional on phenotypic values are estimated directly as variance components and are thus REML estimators.

Although REML estimators of variance components are unbiased, our estimator of $G_{y|x}$ is biased. This bias is comparable to the bias of correlations computed from estimates of variances and covariances. We can derive a rough approximation for the case were both x and y are univariate. We start with computing the expectation of the estimator using an approximation for the expectation of a ratio (Lynch and Walsh '98, p. 818):

$$\begin{aligned} E[G_{y|x}] &= E[G_y] - E[G_{yx}G_x^{-1}G_{yx}] \\ &\simeq E[G_y] - \frac{E[G_{yx}^2]}{E[G_x]}\left(1+CV[G_x]^2 - \frac{Cov[G_{yx}^2, G_x]}{E[G_{yx}^2]E[G_x]}\right), \end{aligned} \tag{6}$$

where CV means coefficient of variation. The expectations pertain to the sampling distribution of the variance components. The covariance in the last term will go to zero as the sampling distribution approaches normality. Ignoring this term, and assuming that the estimates of the variance components are unbiased, gives the following as an asymptotic large-sample expression for the bias

$$\begin{aligned} \text{Bias} &= E[G_{y|x}] - G_{y|x} \\ &\simeq -\frac{Var[G_{yx}]}{E[G_x]}\left(1 + CV[G_x]^2 + \frac{CV[G_x]^2}{CV[G_{yx}]^2}\right). \end{aligned} \tag{7}$$

Thus, ML estimates of conditional variances are on average smaller than the true value. It is possible to use (7) as a bias correction, but a bias-corrected estimator is no longer maximum likelihood, and we will therefore base or discussion on uncorrected estimators. The bias is, however, important to take into account when making general comparisons of variances conditional on genotype with variances conditional on phenotype (which are unbiased) and with pure phenotypic conditional variances (which are less biased due to more precise estimates of the variance compo-

nents) (see Cheverud, '88, '96b for discussion of this point in relation to the comparison of genetic and phenotypic correlations). Equation (7), and a similar consideration of sampling variance (not shown) also show that conditional variances are extremely unreliable if the variance of the constraining variable, x, is close to zero. Since the genetic variance components in our population are all rather small (Hansen et al., 2003), this means that conditioning on genotype is not very accurate. As illustrated in Table 2, the estimated bias is often on the order of 20–30% in our data. For this reason, conditioning on phenotypic values may be more reliable. In general, conditioning on genotypic values with nonsignificant variances will be unreliable and should be avoided.

RESULTS

Evolvabilities

Estimates of I_A for the floral traits are given in Table 1. This shows that if the strengths of selection on these traits are less than 1ϕ (i.e., less than the strength of selection on fitness) during realistic shifts in the selective environment, evolutionary responses are likely to be on the order of a fraction of a percent per generation. Hence, the traits have rather low short-term evolvabilities even when genetic correlations are ignored.

These unconditional evolvabilities are upper limits, as they do not take into account constraints due to correlations among characters. As the genetic correlations given in Appendix A are generally high, we expect the realized evolvability of any subset of traits to be reduced. In Figure 2 we illustrate this by showing the effects on evolvabilities of conditioning on two traits that may often be under stabilizing selection, involucral bract size (UBW) and gland area (GA).

Evolvability conditional on bract size

Due to their important protective function, the involucral bracts may often be under stabilizing selection. The bracts need to fit together and cover the blossom during growth, at night, and finally as the fruit develops. Fruit size in particular may be an important constraining factor (Primack, '87). Bract size may also be a proxy for the overall size of the blossom. To see how much of the evolvability of other blossom traits is tied to this important character, we conditioned on upper-bract width (UBW) as shown in Table 2. We also

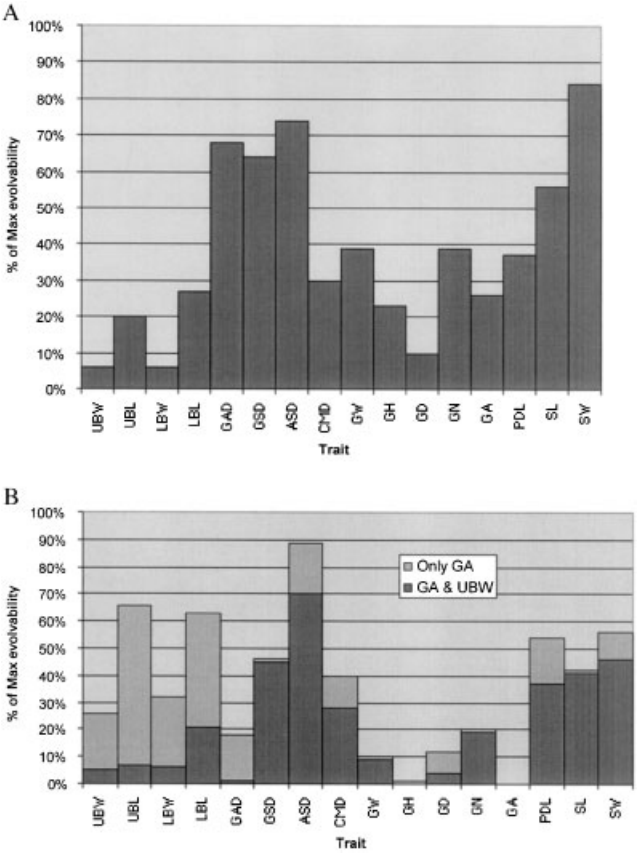

Fig. 2. Conditional evolvability: The columns show the reduction in evolvability, measured as I_A, due to genetic conditioning on A) Upper bract width (UBW) and B) Gland area (GA). In B) the effect of conditioning on GA and UBW simultaneously is also shown.

conditioned on upper and lower bract widths together, obtaining very similar results for non-bract traits (not shown). This indicates that the upper and lower bract have similar influences on the rest of the blossom.

Keeping the bract size phenotypically fixed led to low to moderate reductions in the genetic variance of most non-bract traits. Gland-anther distance (GAD) and style width (SW) even showed an increase in additive genetic variance, but this is most likely due to estimation error. The only traits that seemed to be seriously constrained by keeping bract widths fixed on the phenotypic level were some of the other bract dimensions.

This indicates that the single-generation evolvabilities of most blossom traits are not strongly constrained by bract size. As explained in the theory section, however, it is still possible that this comes about through allowing a (potentially maladaptive) correlated response in bract size.

TABLE 2. Effects of conditioning

	x=UBW[1]				x=GA				x=GAD,GSD[2]		
Trait	Bias	$G_{y\|x}$	$G_{y\|Px}$	h^2	Bias	$G_{y\|x}$	$G_{y\|Px}$	h^2	$G_{y\|x}$	$G_{y\|Px}$	h^2
UBW	−40%	6%	7% ± 3%	0.33->0.11	−23%	26%	69% ± 18%	0.33->0.34	61%	91% ± 22%	0.33->0.46
UBL	−33%	20%	28% ± 9%	0.34->0.26	−15%	66%	90% ± 25%	0.34->0.33	41%	99% ± 25%	0.34->0.38
LBW	−46%	6%	0% ± 2%	0.24->0.00	−27%	32%	63% ± 20%	0.24->0.22	82%	104% ± 27%	0.24->0.38
LBL	−35%	27%	20% ± 7%	0.30->0.19	−18%	63%	85% ± 25%	0.30->0.29	81%	104% ± 27%	0.30->0.40
GAD	−24%	68%	136% ± 57%	0.10->0.15	−27%	18%	79% ± 50%	0.10->0.08	42%	65% ± 46%	0.10->0.07
GSD	−33%	64%	94% ± 23%	0.29->0.41	−26%	46%	87% ± 25%	0.29->0.30	42%	87% ± 26%	0.29->0.27
ASD	−17%	74%	58% ± 20%	0.26->0.18	−11%	89%	87% ± 27%	0.26->0.24	0%	104% ± 29%	0.26->0.29
CMD	−33%	30%	59% ± 45%	0.18->0.12	−24%	40%	51% ± 45%	0.18->0.11	0%	67% ± 45%	0.18->0.13
GW	−26%	39%	79% ± 38%	0.08->0.10	−28%	10%	28% ± 11%	0.08->0.19	0.7%	100% ± 53%	0.08->0.10
GH	−28%	23%	96% ± 36%	0.12->0.16	−29%	0.2%	10% ± 4%	0.12->0.19	35%	75% ± 36%	0.12->0.11
GD	−30%	10%	33% ± 17%	0.44->0.22	−25%	12%	10% ± 10%	0.44->0.13	0%	40% ± 21%	0.44->0.26
GN	−24%	39%	77% ± 40%	0.27->0.21	−25%	20%	20% ± 17%	0.27->0.12	0%	94% ± 45%	0.27->0.26
GA	−27%	26%	83% ± 31%	0.11->0.14	–	–	–	–	18%	84% ± 38%	0.11->0.11
PDL	−30%	37%	40% ± 38%	0.25->0.17	−20%	54%	57% ± 36%	0.25->0.18	46%	78% ± 41%	0.25->0.25
SL	−42%	56%	78% ± 49%	0.12->0.17	−33%	42%	38% ± 50%	0.12->0.09	0%	9% ± 33%	0.12->0.03
SW	−21%	84%	137% ± 29%	0.22->0.42	−23%	56%	116% ± 29%	0.22->0.43	59%	129% ± 29%	0.22->0.45

*For each trait the effects of conditioning on the trait(s) labeled x are shown as percent of the unconditional evolvability (i.e., of I_A, as given in Table 1) that remains after conditioning. The $G_{y|x}$ columns show the effect of conditioning on the genetic component of x, the $G_{y|Px}$ columns show the effects of conditioning on the phenotypic component of x (±SE), and the h^2columns show the change in heritabilities due to phenotypic conditioning. The bias column shows the approximate bias in the maximum-likelihood estimates of $G_{y|x}$ in percent of the unconditional additive genetic variance of the trait. Note that heritabilities, and conditioning on phenotype are based on the larger data set 1 when possible, while conditioning on genotype and all analyses involving CMD, GD, GN, PDL and SL are based on the smaller data set 2.

[1]UBW itself conditioned on LBW.

[2]GAD and GSD conditioned only on each other.

Keeping the genetic value of bract size fixed led to a more severe reduction in evolvability for most traits, suggesting that their evolvability over more than one generation would be significantly reduced by stabilizing selection on bract size or overall blossom size. It should, however, be noted that much of this discrepancy may be due to bias in the estimates.

Evolvability conditional on pollination traits

The gland area (GA), gland-anther distance (GAD) and gland-stigma distance (GSD) are traits with known function in pollination. After a population has adapted to a given selection regime, we expect these traits to be under stabilizing selection. It is thus of interest to assess the extent to which they may constrain the evolution of the other blossom traits. Table 2 shows the effects of conditioning on these traits. Conditioning on the phenotypic value of gland area produced a moderate reduction in genetic variance for many traits, but usually not more than 50% (excepting the other gland traits, which not surprisingly were strongly constrained by GA). The positions of the flowers as measured by GAD and GSD seemed to have less effect on most traits.

Conditioning on the genotypic value of GA reduced evolvability to 40%–70% of the unconditional value for most traits. This is qualitatively similar to the effects of conditioning on bract size, but the differences between phenotypic and genotypic conditioning were less pronounced for GA than for bract size. Keeping GAD and GSD genotypically fixed led to a near complete removal of genetic variation for many traits, but this result is probably invalid due to the rather low levels of genetic variation in GAD and GSD. The conditioning on phenotypic values did not point to strong constraints for most traits.

Architecture of the gland

In addition to assessing evolvability, conditioning can be used to dissect the variational architecture of a character, and estimate its degree of integration. The resin-producing gland is a novel structure, unique to *Dalechampia*, composed of a set of tightly packed bractlets that secrete resin from their edges. The gland area (GA = GH × GW) approximates the area of exposed resin-secreting

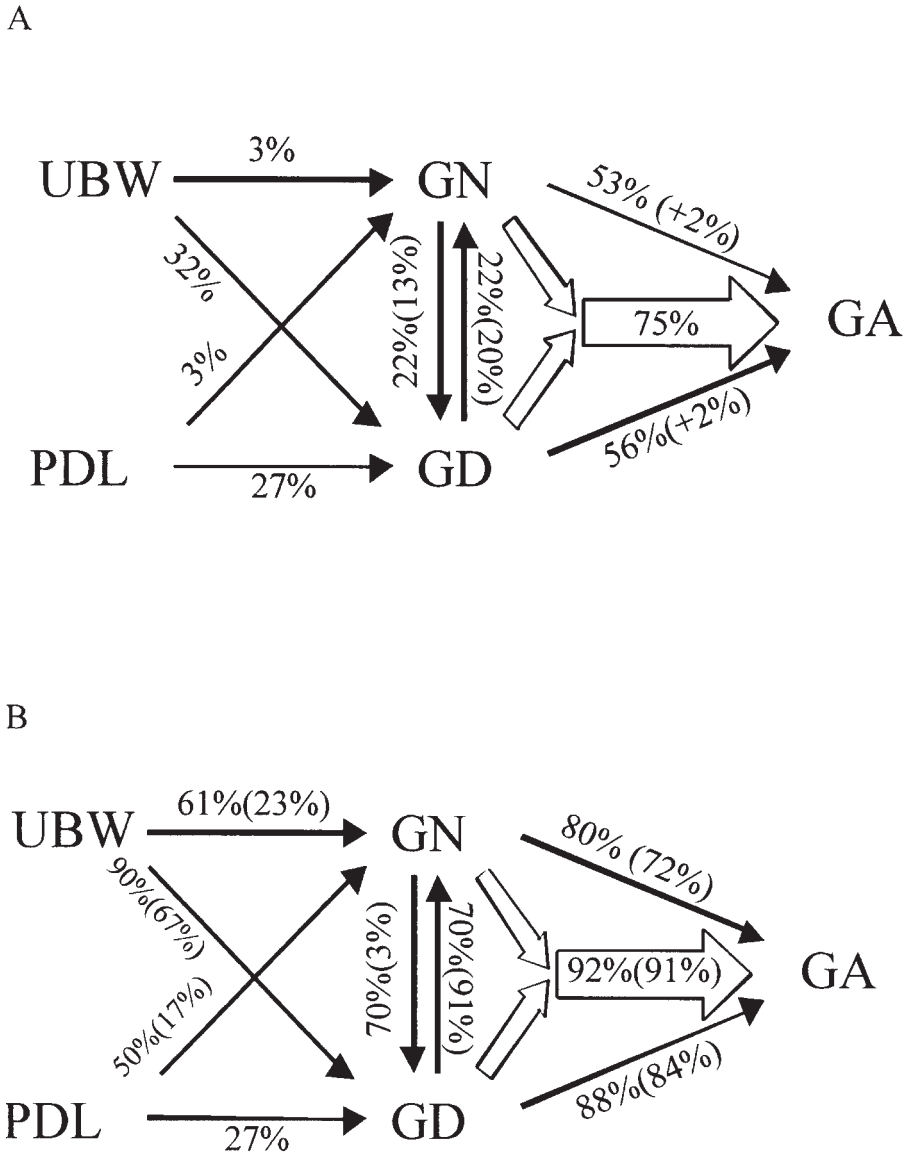

Fig. 3. Variational architecture of the gland: Arrows indicate the amount of variance in the target trait that is explained by the trait at the origin of the arrow. A) Structure of phenotypic variance. B) Structure of additive genetic variance. Each arrow shows % additive genetic variance explained by genetic conditioning and, in parenthesis, % additive genetic variance explained by phenotypic conditioning.

edges and can be considered as a proxy for the average standing crop of resin (Armbruster, '84, but cf. Armbruster and Steiner, '92). The use of resin as a reward for pollinators probably evolved as an exaptation from a floral defense system, where defense resins were produced by bractlets subtending individual male flowers or arms of male flowers (Armbruster et al., '97). The evolution of the gland thus involved extensive reorganization that may have involved migration and duplication of bractlets as well as the elimination of all structures but the bractlets from one or two inflorescence arms. In *D. scandens* the gland consists of two lobes, derived from two arms of male flowers (Froebe and Magin, '93), and each built from 3-4 stacks of bractlets.

The structure of phenotypic variation in the gland is illustrated in Figure 3A. This is based on data set 2 and includes selfed individuals. We first observe that 75% of the variance in GA (broken down as 82% of the variance in GH and 70% of the variance in GW) is explained by the number of bractlets in the gland (GN) and the gland depth (GD). We may think of GD as a proxy for the size

of individual bractlets, as it corresponds closely to the length of the bractlets in the major stacks. Individually, GN and GD explained 53% and 56% of the variation in GA, respectively. Furthermore, the number and size of bractlets are relatively independent. The number of bractlets explained just 22% of variation in GD, and this was reduced to 13–15% when overall size in terms of UBW or PDL was controlled for. In conclusion, the phenotypic variance in GA was largely explained by equal contributions of two moderately independent components, GN and GD, which may be equated with the number and size of bractlets, respectively.

The number of bractlets appears to be phenotypically independent of overall blossom size, as the two size-related traits UBW and PDL each explained only 3% of the variance in GN. The overall blossom size-related variation in GA must then be mediated through bractlet size. This is consistent with UBW and PDL explaining 26% and 22% of the variation in GA, and 32% and 27% of the variation in GD, respectively. Furthermore, adding UBW or PDL to GD and GN as explanatory variables for GA only increased the variance explained by 2%.

We may expect the genetic variation in gland area to be structured in a similar way to the phenotypic variation, with two independent components of which only one is correlated with overall blossom size. It turned out, however, that the genetic component of variation was much more integrated (Fig. 3B). First, both GN and GD individually explained more of the genetic variance in GA. Conditioning GA on GN explained 80% (72% when conditioned on phenotypic value) of the genetic variance, and conditioning on GD explained 88% (84%). Second, conditioning GA on GD and GN combined explained over 92% (91%) of the genetic variation. Third, the genetic effects of GD and GN seemed to be correlated. Genetic conditioning indicated that they share 70% of their genetic variance, but oddly, only 3% of the genetic variance in GD was explained by conditioning on the phenotypic value of GN (61% explained when GN is conditioned on phenotypic value of GD). Finally, the genetic component of GN seems to be dependent on blossom size, as 61% (23%) and 50% (17%) of the genetic variances are accounted for by conditioning on UBW and PDL, respectively.

One interpretation of these results is that, although the gland has two morphologically ''independent'' sources of variation in the size and number of bractlets, these sources of variation may have become integrated on the genetic level. Integration was also supported by very high genetic correlations among the three gland dimensions GW, GH and GD. The correlations remained high after conditioning on UBW, PDL, or GAD (not shown), and are therefore not obviously related to a general blossom-level or cymule-level acquisition effect. Neither were the correlations much reduced when conditioned on bractlet number (not shown), and we conclude that all components of variation seem to constrain the independent evolution of the three gland dimensions.

Evolvability of the gland in relation to interpopulation variation

What is the relationship of these patterns of within-population variation to variation among populations? In Figure 4 we used the genetic model for gland area derived in the previous section to predict variation among five distinct populations. These include a nearby Mexican population (Chetumal), with slightly larger blossoms, and three widely separated Venezuelan populations (Tovar, Caracas, and Puerto Ayacucho) with smaller blossoms adapted to pollination by smaller bees. The success of this prediction suggests that the variation among populations indeed bears resemblance to the genetic architecture of the Tulum population. However, there

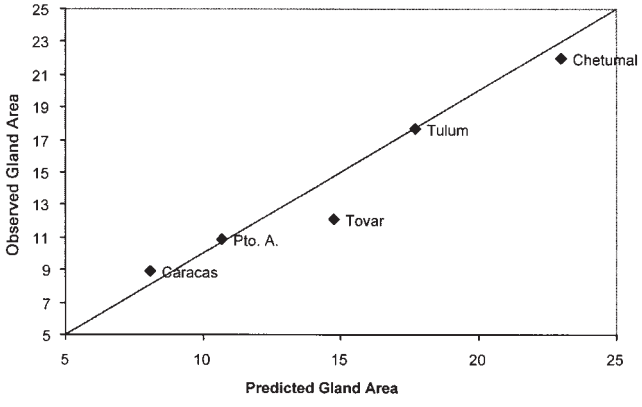

Fig. 4. Observed gland area across populations against gland area predicted from within-population variation: Gland area (GA) is predicted from gland depth (GD) and bractlet number (GN) in the populations according to the equation GA = 11.34GD + 0.49GN, where 11.34 and 0.49 are the partial genetic regressions of GA on GD and GN, respectively, as computed from the Tulum population.

were also aspects of the interpopulation variation that are not predicted from the genetic variation in the Tulum population. The genetic analysis found the number of bractlets in the gland to be evolvable and linked to the majority of the genetic variation in GA. We would therefore predict differences in bractlet number to be the major source of among-population variation in GA. This was not the case. Four of the five populations had roughly similar numbers of gland bractlets. Only in the very small-glanded Caracas population does gland reduction appear to involve loss of bractlets, and even in this population the loss is less than predicted.

If the strong genetic correlations between gland dimensions constitute an evolutionary constraint, populations with different gland sizes should fall on common trajectories predicted by within-population genetic regressions (e.g., Armbruster, '91; Schluter, '96, 2000). Figure 5 shows that differences in gland dimensions among the five populations indeed conform to predictions from the pattern of variation within the Tulum population. One discrepancy is the Tovar population, which seems to have glands that are wider than predicted from their height and depth. Figure 5A also shows that among-population variation in GA is predicted about equally well from GD whether GN is held fixed or not. Holding GN fixed improves the prediction for the nearby Chetumal population, but tends to compromise the prediction for the distant Venezuelan populations. Thus, the gland difference between the two Mexican populations is consistent with a simple change in bractlet size. The differences between Mexican and Venezuelan populations are more complex and may involve changes in size, number, and arrangement of bractlets. This is consistent with the greater genetic distance between the South American populations and the two closely related Mexican populations (unpubl. data).

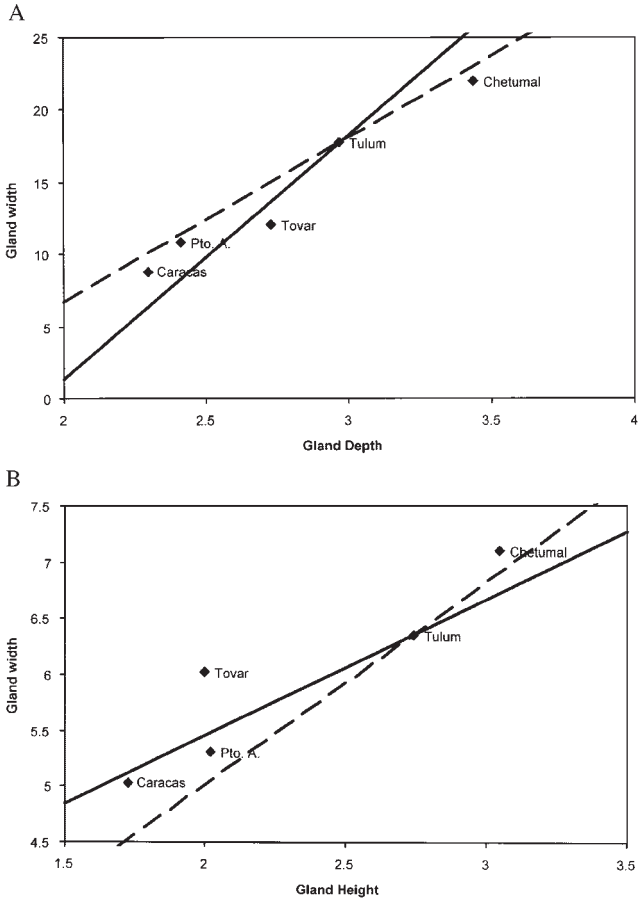

Fig. 5. Predicting interpopulation variation from within-population variation in gland dimensions: The five points are five distinct *Dalechampia scandens* populations. The solid line is the genetic regression based on the Tulum population. The dashed line is the partial genetic regression when number of bractlets (GN) is controlled for. The lines are forced through the position of the Tulum population. (A) Gland area in relation to gland depth, which is a proxy for the size of the bractlets. (B) Gland width in relation to gland height. In both cases the prediction for the more closely related Chetumal population is better when bractlet number is kept fixed, while the prediction for the more distant Venezuelan populations is better when also bractlet number is allowed to vary. The mean bractlet number ($\pm$ SE) for the five popualtions are: Caracas: 17.51 $\pm$ 0.55, Puerto Ayacucho: 20.06 $\pm$ 0.85, Tovar: 21.00 $\pm$ 0.39, Chetumal: 21.39 $\pm$ 0.50, and Tulum: 21.47 $\pm$ 0.24.

DISCUSSION

Not surprisingly, we found that conditional evolvabilities may be less than evolvabilities predicted under the assumption that all the additive genetic variance of a character is available for adaptation. The variance removed by conditioning on phenotypic values of individual characters was small to moderate when the traits were not closely related developmentally. Conditioning on genetic values led to more severe reductions in evolvability, although part of this may be due to bias in the estimates. If taken at face value, the results mean that single-generation evolvabilities are not strongly constrained by other characters individually, but over several generations constraints may be more severe.

These results reinforce the rather low evolvability of the blossoms reported by Hansen et al. (2003). It is noteworthy that this conclusion is reached for traits that seem to be evolutionary

labile at the interpopulation and species level (Armbruster, '85, '88, '93). Although we have been able to explain some of the geographic variation in blossom morphology with selection models, the majority of variation is still unaccounted for (Armbruster, '90; Hansen et al., 2000). This may be due to geographic variation in selective factors not included in our crude selection models or it may be a reflection of genetic constraints. The low evolvabilities suggest that local populations may be unable to track rapid environmental changes, such as may occur with changes in the composition of bee communities or with changes in the relative abundance of competing *Dalechampia* species.

We expect the conditional genetic variance to decrease even further as more characters are included in the conditioning, and the evolvability may in reality be determined by the genetic variation that remains when all functionally unrelated traits are held fixed. In our data the rather low levels of genetic variation found at the outset made it unfeasible to condition on many variables at once, and it remains an open question how much further the evolvability would be reduced. Several studies have shown that a large amount of multivariate genetic variation within and among species can be explained by very few degrees of freedom (Björklund, '96; Schluter, 2000, chap. 9; Houle, 2001). This suggests that the independent evolvability of single characters may often be very low, and selection may be forced to act on suites of characters along genetic lines of least resistance (Schluter, '96). If true, this means that genetic constraints may be crucial in understanding evolutionary change, and explains why patterns of quantitative genetic variation sometimes do resemble among-species variation (e.g., Mitchell-Olds, '96; Schluter, '96; Andersson, '97; Merilä and Björklund, '99; this study; but see Herrera et al., 2002 for a botanical counter example).

Wagner ('96) suggested that modularity might evolve through a mosaic combination of stabilizing and recurring directional selection. This may favor the suppression of genes with pleiotropic effects on characters that do not tend to be selected in concert and an enhancement of genes with pleiotropic effects on functionally related characters. The result is a parcellation of characters along functional lines that may enhance the evolvability of the organism. Diggle ('92) and Conner and Via ('93) provide some possible examples of functional integration in floral evolution that may be interpreted in this light. As there seems to have been repeated shifts between a few well-defined adaptive regimes, *D. scandens* blossoms are good candidates for this type of selection. We may thus predict that functionally related traits such as GA, GAD, and GSD should become genetically integrated (Armbruster, '91; Armbruster and Schwaegerle, '96). The strong genetic correlations among these morphologically rather separate traits (which were not much reduced by conditioning on other traits), are consistent with this hypothesis, but the rather low levels of genetic variation in GAD and GSD make this a tentative result. The genetic architecture of the gland itself provides stronger evidence for Wagner's hypothesis. At the morphological level, gland variation seems to have two largely independent sources, which we interpreted as variation in the size and number of bractlets. On the genetic level, however, these two sources appears to be more integrated, as the Wagner hypothesis predicts if there has been shifting directional selection on gland size, as ecological data suggest has been the case (Armbruster, '85; Hansen et al., 2000).

One trait that was severely affected by conditioning was style length, SL, conditioned on gland-stigma distance, GSD. Although this is partly a consequence of both traits being affected by the position of the style tip, it is nevertheless a biologically significant constraint. Sexual selection in *Dalechampia*, as manifested through pollen competition, should select for longer styles (Armbruster, '96, 2001). This is because the large stigmatic surface, which extends more than halfway down the style in *D. scandens* creates the potential for huge variance in the distance that pollen tubes grow to reach the ovules, thus compromising the effectiveness of ''genetic screening'' though pollen competition (Armbruster et al., '95). The only way to ''correct'' this is to confine the stigmatic surface to the style tip (involving dramatic expansion of the tip if stigmatic area is to remain constant) or increase the distance tubes must grow. While the former has evolved in a few species, it appears to be of limited evolvability in *D. scandens* and its relatives (Armbruster et al., '95). The even more severe constraint on the evolution of style length imposed by stabilizing selection on GSD may have led to the unusual ''solution'' of increasing the length of pollen-tube growth without changing the stigmatic fit with pollinators: bidirectional pollen-tube growth in pollen landing on the stigmatic surfaces on the sides of the style (Armbruster et al., '95; Armbruster, '96).

Zhu (’95) was the first to suggest explicitly that genetic variation could be conditioned on the phenotypic values of other characters, and this was applied to study late growth independently of early growth in lab mice (Atchley and Zhu, ’97). An interesting botanical application is provided by the recent work of Worley and Barrett (2000), who studied the trade-off between flower size and number in the emergent aquatic plant *Eichhornia paniculata*. They estimated the change in additive genetic variances and correlations due to holding fixed two phenotypic variables related to inflorescence size. They found a 40% decrease in the additive genetic variance of flower size and a 10% decrease for flower number. They also found a minor change in the genetic correlation between these two traits from 0.15–0.18 to −0.07. This was interpreted as removing some common acquisition variation at the level of the inflorescence.

Of course, any study that corrects for size or related variables before computing genetic variance components employs phenotypic conditioning, but the reduction in genetic variation has rarely been interpreted as a measure of evolutionary constraint. This may be because the common, but in our view mistaken, use of heritability as a measure of evolvability (see Houle, ’92; Hansen et al., 2003) will obscure the effects of conditioning, because the reduction in genetic and phenotypic variances will tend to cancel. In our data, heritabilities sometimes increased, sometimes decreased, and were sometimes left unaffected by conditioning (Table 2).

Conditioning is essentially a tool for studying variance matrices and can be used along with techniques such as common principal components, factor analysis or other methods (e.g., Wagner ’84; Zelditch, ’88; Cheverud et al., ’89; Schluter, ’96; Steppan, ’97; Steppan et al., 2002) to dissect the variational structure of a character. There is no particular statistical advantage to conditioning, but there are conceptual benefits. We have demonstrated that conditional variances have a clear theoretical relation to evolvability, and may therefore be useful when studies focus on evolvability or genetic constraints. Conditioning is also a natural tool for the study of modularity, as it allows assessment of the relative independence of character complexes (see Olson and Miller, ’58; Cheverud, ’96a; Mezey et al., 2000; Magwene, 2001 for related ideas).

The extent to which genetic variation can be used to infer underlying genetic structure is an open question. Linking genetic correlation to functional organization of the organism has been criticized; especially in relation to detecting negative pleiotropy in the face of variation in general vigor genes (Charlesworth, ’90; Houle, ’91; de Jong and van Noordwijk, ’92; Fry, ’93). It is also clear that different sorts of pleiotropy may cancel and produce little genetic correlation on average (e.g., Baatz and Wagner, ’97). Although patterns of pleiotropy can thus not be inferred directly from genetic correlations, it may still be possible to test hypotheses about modularity against estimates of the amount of variation that is left after conditioning on other characters. After all, if a character is to be considered a quasi-independent module, there must exist some genetic variation that is independent of the rest of the organism. For example, Berg’s (’60) hypothesis that floral and vegetative traits constitute independent modules, or ‘‘correlation pleiades,’’ in plants could be investigated by conditioning floral on vegetative traits and vice versa (see Conner and Sterling, ’95, ’96; Waitt and Levin, ’98; Armbruster et al., ’99; Magwene, 2001; Herrera et al., 2002).

It may also be objected that segregating genetic variation is an ephemeral aspect of evolvability, important for short-term adjustments to minor variation in adaptive optima, but perhaps less relevant for long-term adaptation to significant changes in the environment. One aspect of this problem is that genetic variances and covariances are not merely a reflection of variability, but are also directly influenced by stabilizing selection acting on the character (Wagner, ’89; Arnold, ’92; Wagner et al., ’97). Thus, the ability of the genetic system to produce variation is confounded with the effects of selection on that variation. This suggests that studies of evolvability should rather focus on mutational variation, which may provide a more direct window into the genotype-phenotype map than the study of segregating variation.

On the other hand, the estimation of mutational variance matrices is notoriously difficult (see e.g., Camara and Pigliucci, ’99 for an attempt with *Arabidopsis*). We also note that an unknown quantity of new mutations may be unavailable for adaptation due to deleterious pleiotropic side effects. Mutations in general housekeeping genes or in the coding region of genes expressed in a multitude of tissues or circumstances may introduce variation in a given phenotypic character, but this variation cannot be effectively utilized by selection on the character. Galis (’99) suggests an intriguing example in which apparent genetic

variability in the number of mammalian neck vertebrae is rendered useless by pleiotropic effects on cancer risk. In other words, evolvability is constrained by internal selection (see Galis and Metz, 2001; Galis et al., 2001 for further examples). We want to eliminate such constrained variation from a measure of character evolvability. It may be possible to address this problem by conditioning on a measure of background fitness independent of trait in question; but if this is not practical, standing genetic variation has the advantage, relative to mutational variation, of being somewhat pruned of this confounding component of variation.

A final problem is that genetic variation is likely to change during a prolonged response to selection (e.g., Turelli, '88; Arnold, '92). This is partially for the simple reason that alleles may go to fixation, but more fundamentally because the genetic variability, in the proper sense of ability to vary, will change, because the genotype-phenotype map itself will be changing under directional selection unless gene interactions are truly additive on a functional level (Hansen and Wagner, 2001). Both gene effects and mutational variation will change in ways that depend on the functional epistatic interactions among genes. The evolutionary significance of measures of short-term evolvability is perhaps best regarded as an empirical question that may eventually be answered through testing the predictions against patterns of evolutionary change on multiple levels. But regardless of level, the evolvability of quantitative characters needs to be reassessed with an eye to how much of the genetic variation can reasonably be expected to be useful for adaptation.

ACKNOWLEDGMENTS

We thank David Houle, Jason Mezey, Martin Morgan, Günter Wagner and the anonymous reviewers for discussions and comments on the manuscript, and many people at NTNU for help in the greenhouse.

LITERATURE CITED

Anderson TW. 1984. An introduction to multivariate statistical analysis. Sec. ed., Wiley & Sons. New York.

Andersson S. 1997. Genetic constraints on phenotypic evolution in *Nigella* (Ranunculaceae). Bio J Linn Soc 62:519–532.

Armbruster WS. 1984. The role of resin in angiosperm pollination: ecological and chemical considerations. Am J Bot 71:1149–1160.

Armbruster WS. 1985. Patterns of character divergence and the evolution of reproductive ecotypes of *Dalechampia scandens* (Euphorbiaceae). Evolution 39:733–752.

Armbruster WS. 1988. Multilevel comparative analysis of the morphology, function, and evolution of *Dalechampia* blossoms. Ecology 69:1746–1761.

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

Armbruster WS. 1991. Multilevel analyses of morphometric data from natural plant populations: insights into ontogenetic, genetic, and selective correlations in *Dalechampia scandens*. Evolution 45:1229–1244.

Armbruster WS. 1993. Evolution of plant pollination systems: hypotheses and tests with the Neotropical vine *Dalechampia*. Evolution 47:1480–1505.

Armbruster WS. 1996. Evolution of floral morphology and function: An integrative approach to adaptation, constraint, and compromise in *Dalechampia* (Euphorbiaceae). In Lloyd DG, Barrett SCH (eds), Floral biology: studies on floral evolution in animal-pollinated plants, pp. 241–272. Chapman & Hall, New York.

Armbruster WS. 2001. Evolution of floral form: Electrostatic forces, pollination, and adaptive compromise. New Phytologist 152:181–183.

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

Armbruster WS, Steiner KE. 1992. Pollination ecology of four *Dalechampia* species (Euphorbiaceae) in northern Natal, South Africa. Am J Bot 79:306–313

Armbruster WS, Martin P, Kidd J, Stafford R, Rogers DG. 1995. Reproductive significance of indirect pollen-tube growth in *Dalechampia* (Euphorbiaceae). Am J Bot 82:51–56.

Armbruster WS, Howard JJ, Clausen TP, Debevec EM, Loquvam JC, Matsuki M, Cerendolo B, Andel F. 1997. Do biochemical exaptation link evolution of plant defence and pollination systems? historical hypotheses and experimental tests with *Dalechampia* vines. Am Nat 149:461–484.

Armbruster WS, Di Stilio VS, Tuxill JD, Flores TC, Velásquez Runke JL. 1999. Covariance and decoupling of floral and vegetative traits in nine Neotropical plants: a re-evaluation of Berg's correlation-pleiades concept. Am J Bot 86:39–55.

Arnold S. 1992. Constraints on phenotypic evolution. Am Nat 140:S85–S107.

Atchley WR, Zhu J. 1997. Developmental quantitative genetics, conditional epigenetic variability and growth in mice. Genetics 147:765–776.

Baatz M, Wagner GP. 1997. Adaptive inertia caused by hidden pleiotropic effects. Theoretical Population Biology 51:49–65.

Berg R. 1960. The ecological significance of correlation pleiades. Evolution 17: 171–180.

Björklund M. 1996. The importance of evolutionary constraints in ecological time scales. Evolutionary Ecology 10:423–431.

Camara MD, Pigliucci M. 1999. Mutational contributions to genetic variance/covariance matrices: an experimental approach using induced mutations in *Arabidopsis thaliana*. Evolution 53:1692–1703.

Charlesworth B. 1990. Optimization models, quantitative genetics, and mutation. Evolution 44:520–538.

Cheverud JM. 1988. A comparision of genetic and phenotypic correlations. Evolution 42: 958–968.

Cheverud JM. 1996a. Developmental integration and the evolution of pleiotropy. Amer Zool 36:44–50.

Cheverud J. 1996b. Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. J Evol Biol 9:5–42.

Cheverud JM. 2001. The genetic architecture of pleiotropic relations and differential epistasis. pp 411–433, In: Wagner GP (ed), The character concept in evolutionary biology. Academic Press, SanDiego.

Cheverud JM, Wagner GP, Dow MM. 1989. Methods for the comparative analysis of variation patterns. Syst Zool 38:201–213.

Conner JK, Sterling A. 1995. Testing hypotheses of functional relationships: a comparative survey of correlation patterns among floral traits in five insect-pollinated plants. Am J Bot 82:1399–1406.

Conner JK, Sterling A. 1996. Selection for independence of floral and vegetative traits: evidence from correlation patterns in five species. Can J Bot 74:642–644.

Conner JK, Via S. 1993. Patterns of phenotypic and genetic correlations in wild radish: Possible evidence for the effects of selection on floral morphology. Evolution 47: 704–711.

De Jong G, van Noordwijk AJ. 1992. Acquisition and allocation of resources: genetic (co)variances, selection, and life histories. Am Nat 139:749–770.

Diggle PK. 1992. Development and the evolution of plant reproductive characters. In R. Wyatt (ed.), Ecology and evolution of plant reproduction: New approaches, pp. 326–355. Chapman & Hall, New York.

Froebe HA, Magin N. 1993. Pattern analysis in the inflorescence of *Dalechampia* L. (Euphorbiaceae). Bot Jahrb Syst 115:27–44.

Fry JD. 1993. The ''general vigor'' problem: can antagonistic pleiotropy be detected when genetic covariances are positive? Evolution 47:327–333.

Galis F. 1999. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. J Exp Zool (Mol Dev Evol) 285: 19–26.

Galis F, Metz JAJ. 2001. Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation. J Exp Zool (Mol Dev Evol) 291:195–204.

Galis F, van Alphen JM, Metz JAJ. 2001. Why five fingers? Evolutionary constraint on digit numbers. TREE 16: 637–646.

Gerhart J, Kirschner M. 1997. Cells, embryos and evolution: towards a cellular and developmental understanding of phenotypic variation and evolutionary adaptability. Blackwell. USA.

Hansen TF. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Hansen TF. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. BioSystems (in press).

Hansen TF, Boonstra R. 2000. The best in all possible worlds? A quantitative genetic study of geographic variation in the Meadow vole (*Microtus pennsylvanicus*). Oikos: 89: 81–94.

Hansen TF, Wagner GP. 2001. Modeling genetic architecture: A multilinear model of gene interaction. Theor Pop Biol 59:61–86.

Hansen TF, Armbruster WS, Antonsen L. 2000. Comparative analysis of character displacement and spatial adaptations as illustrated by the evolution of *Dalechampia* blossoms. Am Nat 156:S17–S34.

Hansen TF, Pélabon C, Armbruster WS, Carlson ML. 2003. Evolvability and constraint in *Dalechampia* blossoms: Components of variance and measures of evolvability. J Evol Biol (in press).

Herrera CM, Cerda X, Garcia MB, Guitian J, Medrano M, Rey PJ, Sanchez-Lafuente AM. 2002. Floral integration, phenotypic covariance structure and pollinator variation in bumblebee-pollinated *Helleborus foetidus*. J Evol Biol 15:108–121.

Houle D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630–648.

Houle D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–204.

Houle D. 2001. Characters as the units of evolutionary change. In Wagner GP, ed. The character concept in evolutionary biology. pp. 109–140. Academic press. San Diego.

Kauffman SA. 1993. The origins of order. Self-organization and selection in evolution. Oxford Univ Press. Oxford.

Kondrashov AS, Turelli M. 1992. Deleterious mutations, apparent stabilizing selection and the maintainance of quantitative variation. Genetics 132:603–618.

Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33:402–416.

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

Lewontin RC. 1978. Adaptation. Scient Am 239:212–231.

Lynch M, Walsh B. 1998. Genetics and analysis of quantitative characters. Sinauer. Sunderland, Massachussets.

Magwene PM. 2001. New tools for studying integration and modularity. Evolution 55: 1734–1745.

Merilä, J, Björklund, M. 1999. Population divergence and morphometric integration in the greenfinch (*Carduelis chloris*)—evolution against the trajectory of least resistance? J Evol Biol 12:103–112.

Mezey JG, Cheverud JM, Wagner GP. 2000. Is the genotype-phenotype map modular?: a statistical approach using mouse quantitative trait loci. Genetics 156:305–311.

Mitchell-Olds T. 1996. Pleiotropy causes long-term genetic constraints on life-history evolution in *Brassica rapa*. Evolution 50:1849–1858.

Olson EC, Miller RL. 1958. Morphological integration. Chicago Univ. Press. Chicago.

Primack RB. 1987. Relationships among flowers, fruits, and seeds. Ann Rev Ecol Syst 18:409–430.

Raff RA. 1996. The shape of life: genes, development, and the evolution of animal form. Univ. Chicago press. Chicago.

Riedl RJ. 1977. A systems-analytical approach to macroevolutionary phenomena. Quart Rev Biol 52:351–370.

Riedl RJ. 1978. Order in living organisms: a systems analysis of evolution. Wiley, New York.

Rowe L, Houle D. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. Proc R Soc B 263:1415–1421.

Schluter D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50: 1766–1774.

Schluter D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

Searle SR, Casella G, McCulloch CE. 1992. Variance components. Wiley, New York.

Steppan SJ. 1997. Phylogenetic analysis of phenotypic covariance structure. I. contrasting results from matrix

correlation and common principal component analyses. Evolution 51:571–586.
Steppan SJ, Phillips PC, Houle D. 2002. Comparative quantitative genetics: evolution of the G matrix. TREE 17:320–327.
Stern D. 2000. Perspective: Evolutionary developmental biology and the problem of variation. Evolution 54: 1079–1091.
Turelli M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42:1342–1347.
Von Dassow G, Munro E. 1999. Modularity in animal development and evolution: elements of a conceptual framework for EvoDevo. J. Exper. Zool. (MDE): 285: 307–325.
Wagner GP. 1984. On the eigenvalue distribution of genetic and phenotypic dispersion matrices: Evidence for a nonrandom organization of quantitative character variation. J Math Biol 21:77–95.
Wagner GP. 1989. Multivariate mutation-selection balance with constrained pleiotropic effects. Genetics 122:223–234.
Wagner GP. 1996. Homologues, natural kinds and the evolution of modularity. Am Zool 36:36–43.
Wagner GP, Altenberg L. 1996. Complex adaptations and evolution of evolvability. Evolution 50:967–976.
Wagner GP, Booth G, Bagheri-Chaichian H. 1997. A population genetic theory of canalization. Evolution 51:329–347.
Waitt DE, Levin DA. 1998. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. Heredity 80:310–319
Williams GC. 1992. Natural selection: Domains, levels, and challenges. Oxford Univ. Press. Oxford.
Worley AC, Barrett SCH. 2000. Evolution of floral display in *Eichhornia paniculata* (Pontederiaceae): Direct and correlated responses to selection on flower size and number. Evolution 54:1533–1545.
Zelditch ML. 1988. Ontogenetic variation in pattens of phenotypic integration in the laboratory rat. Evolution 42:28–41.
Zhu J. 1995. Analysis of conditional genetic effects and variance components in developmental genetics. Genetics 141:1633–1639.

APPENDIX A: The G-matrix: Above the diagonal are additive genetic covariances (± SE), and below the diagonal are additive genetic correlations and, in bold, phenotypic correlations. To obtain the covariances each pair of traits is analyzed separately with a model that includes an additive genetic and an environmental covariance matrix. The genetic variances given on the diagonal are computed separately based on a model including only an additive genetic and a residual variance component (these are usually almost identical to the estimates from the pairwise analyses)

	UBW	UBL	LBW	LBL	GAD	GSD	ASD	CMD	GW	GH	GD	GN	GA	PDL	SL	SW
UBW	1.86	1.25	1.68	1.36	0.19	0.23	0.41	0.12	0.44	0.36	0.22	2.47	3.60	0.34	0.26	.046
	±.79	±.55	±.86	±.68	±.13	±.18	±.26	±.060	±.20	±.14	±.080	±1.05	±1.43	±.17	±.21	±.0.45
UBL	0.88	1.05	1.21	1.11	0.17	.079	0.39	.068	0.21	0.19	0.13	1.44	1.83	0.30	0.12	−006
	0.82	±.43	±.60	±.51	±.10	±.12	±.20	±.041	±.14	±.10	±.060	±.76	±.98	±.13	±.15	±.0.30
LBW	0.97	0.93	1.62	1.37	0.10	0.15	0.49	0.11	0.36	0.34	0.21	2.09	3.21	0.30	0.17	.031
	0.91	**0.74**	±1.01	±.78	±.14	±0.20	±.28	±.070	±.22	±.16	±.090	±1.17	±1.57	±.19	±.23	±.051
LBL	0.85	0.93	0.91	1.36	0.12	.079	0.50	.054	0.25	0.22	0.16	1.56	2.18	0.31	0.12	.008
	0.82	**0.89**	**0.86**	±.68	±.12	±.15	±.24	±.051	±.17	±.12	±.070	±.94	±1.23	±.16	±.19	±.039
GAD	0.59	0.63	0.32	0.41	.060	.052	.075	.029	0.10	.058	.036	0.51	0.68	.052	.074	.003
	0.31	**0.25**	**0.23**	**0.15**	±.037	±.035	±.052	±.012	±.040	±.028	±.016	±.23	±.28	±.032	±.041	±.009
GSD	0.60	0.28	0.44	0.25	0.87	.078	-.028	.029	.095	.060	.047	0.20	0.63	.029	.066	.012
	0.69	**0.45**	**0.68**	**0.52**	**0.29**	±.055	±.064	±.016	±.050	±.035	±.022	±.26	±.35	±.040	±.057	±.012
ASD	0.51	0.66	0.69	0.78	0.50	−0.17	0.35	.012	.069	.054	.013	0.74	0.61	0.14	−012	−007
	0.33	**0.53**	**0.22**	**0.30**	**0.02**	**0.16**	±.15	±.023	±.074	±.052	±.029	±.44	±.52	±0.07	±.075	±.017
CMD	0.78	0.60	0.77	0.42	0.96	0.87	0.19	.011	.027	.027	.018	0.25	0.25	.013	.022	.005
	0.36	**0.27**	**0.32**	**0.09**	**0.06**	**0.15**	**0.16**	±.007	±.017	±.013	±.008	±.09	±.12	±.014	±.018	±.004
GW	0.77	0.49	0.67	0.51	1.04	0.85	0.28	0.62	0.17	0.11	.066	0.74	1.20	.065	.099	.027
	0.51	**0.35**	**0.45**	**0.35**	**0.22**	**0.49**	**0.17**	**0.42**	±0.07	±.040	±.026	±.35	±.47	±.047	±.058	±.014
GH	0.86	0.59	0.82	0.60	0.79	0.74	0.30	0.81	0.93	.090	.049	0.65	0.92	.069	.063	.014
	0.50	**0.37**	**0.43**	**0.36**	**0.23**	**0.48**	**0.16**	**0.41**	**0.76**	±.035	±.018	±.26	±.34	±.034	±.041	±.009
GD	0.87	0.70	0.86	0.72	0.88	0.93	0.12	0.89	0.92	0.95	.029	0.33	0.49	.042	.056	.009
	0.58	**0.48**	**0.51**	**0.46**	**0.29**	**0.52**	**0.44**	**0.48**	**0.77**	**0.73**	±.011	±.13	±.18	±.020	±.026	±.005
GN	0.77	0.57	0.67	0.53	0.85	0.31	0.51	1.10	0.79	0.94	0.86	5.36	6.36	0.51	0.24	.090
	0.20	**0.08**	**0.16**	**0.03**	**0.12**	**0.20**	**0.09**	**0.25**	**0.68**	**0.71**	**0.51**	±2.36	±2.57	±0.26	±0.30	±.069
GA	0.83	0.56	0.78	0.58	0.89	0.75	0.33	0.75	0.97	0.99	0.94	0.90	9.42	0.65	0.67	0.17
	0.53	**0.39**	**0.46**	**0.37**	**0.24**	**0.50**	**0.18**	**0.43**	**0.91**	**0.96**	**0.78**	**0.74**	±3.46	±0.34	±0.41	±.10
PDL	0.79	0.89	0.76	0.82	0.68	0.33	0.76	0.41	0.51	0.72	0.75	0.68	0.67	.098	.062	-.002
	0.63	**0.65**	**0.57**	**0.65**	**0.30**	**0.42**	**0.40**	**0.29**	**0.47**	**0.45**	**0.52**	**0.16**	**0.47**	±.050	±.051	±.010
SL	0.68	0.43	0.48	0.38	1.14	0.85	−0.07	0.67	1.04	0.83	1.14	0.37	0.91	0.70	.082	.012
	0.66	**0.57**	**0.63**	**0.62**	**0.26**	**0.75**	**0.06**	**0.33**	**0.63**	**0.64**	**0.65**	**0.28**	**0.67**	**0.59**	±.077	±.015
SW	0.42	−.08	0.32	0.09	0.15	0.53	−0.15	0.53	0.83	0.55	0.63	0.48	0.67	−0.10	0.53	.007
	0.56	**0.46**	**0.53**	**0.52**	**0.25**	**0.60**	**0.07**	**0.32**	**0.61**	**0.58**	**0.55**	**0.24**	**0.61**	**0.52**	**0.82**	±.004