

## Plasticity of Size and Growth in Fluctuating Thermal Environments: Comparing Reaction Norms and Performance Curves<sup>1</sup>

JOEL G. KINGSOLVER,<sup>2,\*</sup> RIMA IZEM,<sup>3,†</sup> AND GREGORY J. RAGLAND\*

*\*Department of Biology and*

*†Department of Statistics, University of North Carolina, Chapel Hill NC 27599*

**SYNOPSIS.** Ectothermic animals exhibit two distinct kinds of plasticity in response to temperature: Thermal performance curves (TPCs), in which an individual's performance (*e.g.*, growth rate) varies in response to current temperature; and developmental reaction norms (DRNs), in which the trait value (*e.g.*, adult body size or development time) of a genotype varies in response to developmental temperatures experienced over some time period during development. Here we explore patterns of genetic variation and selection on TPCs and DRNs for insects in fluctuating thermal environments. First, we describe two statistical methods for partitioning total genetic variation into variation for overall size or performance and variation in plasticity, and apply these methods to available datasets on DRNs and TPCs for insect growth and size. Our results indicate that for the datasets we considered, genetic variation in plasticity represents a larger proportion of the total genetic variation in TPCs compared to DRNs, for the available datasets. Simulations suggest that estimates of the genetic variation in plasticity are strongly affected by the number and range of temperatures considered, and by the degree of nonlinearity in the TPC or DRN. Second, we review a recent analysis of field selection studies which indicates that directional selection favoring increased overall size is common in many systems—that bigger is frequently fitter. Third, we use a recent theoretical model to examine how selection on thermal performance curves relates to environmental temperatures during selection. The model predicts that if selection acts primarily on adult size or development time, then selection on thermal performance curves for larval growth or development rates is directly related to the frequency distribution of temperatures experienced during larval development. Using data on caterpillar temperatures in the field, we show that the strength of directional selection on growth rate is predicted to be greater at the modal (most frequent) temperatures, not at the mean temperature or at temperatures at which growth rate is maximized. Our results illustrate some of the differences in genetic architecture and patterns of selection between thermal performance curves and developmental reaction norms.

### INTRODUCTION

Temperature has two distinct effects on the body size of an ectothermic organism. First, temperature impacts the growth and developmental rates of an organism, altering the time or age at which a particular body size is achieved (Angilletta *et al.*, 2003). Second, in many ectotherms temperature conditions experienced during development may affect final (adult) body size (Atkinson, 1993; Atkinson and Sibly, 1997). Most of the contributions to this symposium address one or both of these consequences of temperature in a variety of systems.

If we consider growth rate, development time or adult size as a quantitative trait  $z$ , then the value of that trait for a genotype varies as a function of temperature  $T$ : that is, the trait is a curve or function,  $z(T)$ . The effects of temperature on the expression of such traits may thus be viewed as examples of phenotypic plasticity: the phenotypic trait expressed by a genotype depends on environment (Stearns, 1989). In this paper we will distinguish between two different types of plasticity: the thermal developmental reaction norm

(TDRN), in which temperature during growth and development affects the value of some trait expressed by a genotype at some later stage (*e.g.*, adult body mass or time to maturity) (Scheiner, 1993); and the thermal performance curve (TPC), in which current temperature affects the level of some aspect of performance (*e.g.*, growth or development rate) in an individual (Huey and Stevenson, 1979).

Both developmental reaction norms and performance curves are forms of phenotypic plasticity, so that we can use a similar quantitative genetic framework to explore phenotypic and genetic variation, selection and microevolution of these traits (de Jong and Bijma, 2002; Gomulkiewicz and Kirkpatrick, 1992; Via and Lande, 1985). However, thermal developmental reaction norms (TDRNs) and performance curves (TPCs) differ in several important ways. First, for a TDRN an individual only expresses a single value of the trait (adult body mass, development time) during its lifetime, depending on the particular thermal conditions it experienced during its development. By contrast, as the body temperature of an individual changes, the individual's performance may also change, so that an individual may express a wide range of values of the trait (*e.g.*, growth or development rate) during its lifetime. In this sense a TDRN is a property of a genotype, whereas a TPC is a property of an individual. Second, developmental reaction norms are typically irreversible: changing or reversing environmental con-

<sup>1</sup> From the Symposium *Evolution of Thermal Reaction Norms for Growth Rate and Body Size in Ectotherms* presented at the Annual Meeting of the Society for Integrative and Comparative Biology, 5–9 January 2004, at New Orleans, Louisiana.

<sup>2</sup> E-mail: jgking@bio.unc.edu

<sup>3</sup> Current address: Department of Statistics, Harvard University, Cambridge MA 02138.

ditions after some critical stage in development will not alter the expressed phenotypic trait of that individual. By contrast the performance of an individual may change rapidly and reversibly with changes in environmental conditions. Third, developmental reaction norms and performance curves typically operate at different time scales. For example, one can consider the short-term or instantaneous rate of growth rate of an individual, whereas growth to adult size is defined over some much longer period of development. In constant environmental conditions these differences are largely irrelevant. In most terrestrial and many aquatic environments, however, temperature varies substantially over hourly, daily and seasonal time scales that impact growth and developmental rates. We note that a TPC may also exhibit thermal acclimation, such that the shape or position of the TPC is irreversibly affected by developmental temperatures: in this sense the position of the TPC may also be a TDRN (Huey and Berrigan, 1996; Huey *et al.*, 1999).

Here we explore some of the similarities and differences between thermal performance curves and developmental reaction norms, and how these may affect variation, selection and evolution of the thermal plasticity of growth and size in insects. First, we describe a quantitative method for decomposing genetic variation of TDRNs and TPCs into variation due to overall size or performance across all temperatures and variation in the thermal sensitivity of size or performance at different temperatures. We apply this method to a series of data sets for TDRNs and TPCs of adult size, developmental time and performance for insects. Second, we review a recent analysis of studies of selection on size in a natural population, to explore whether patterns of selection on size and development time are different than selection on other phenotypic traits. Third, we describe a recent theoretical model that predicts the relationship between variation in the thermal environment and directional selection on thermal performance curves for growth rate. We use this model to examine how we might distinguish selection on adult size from selection on TPCs for growth rate. Our results highlight some of the important differences between developmental reaction norms and performance curves in terms of both genetic architecture and patterns of selection.

#### GENETIC VARIATION IN REACTION NORMS AND PERFORMANCE CURVES

Many studies have now examined genetic variation in thermal plasticity of size, performance, developmental time or other traits in ectotherms (David *et al.*, 1994; Gilchrist, 1996; Scheiner, 1993; Schlichting and Pigliucci, 1998). Typically, individuals from a clone, family, genetic line or population are reared at several different (constant or fluctuating) temperatures. For a developmental reaction norm (DRN), different individuals (or ramets) from each family or clone are reared at different developmental temperatures; the mean phenotypic value (*e.g.*, adult size, development time,

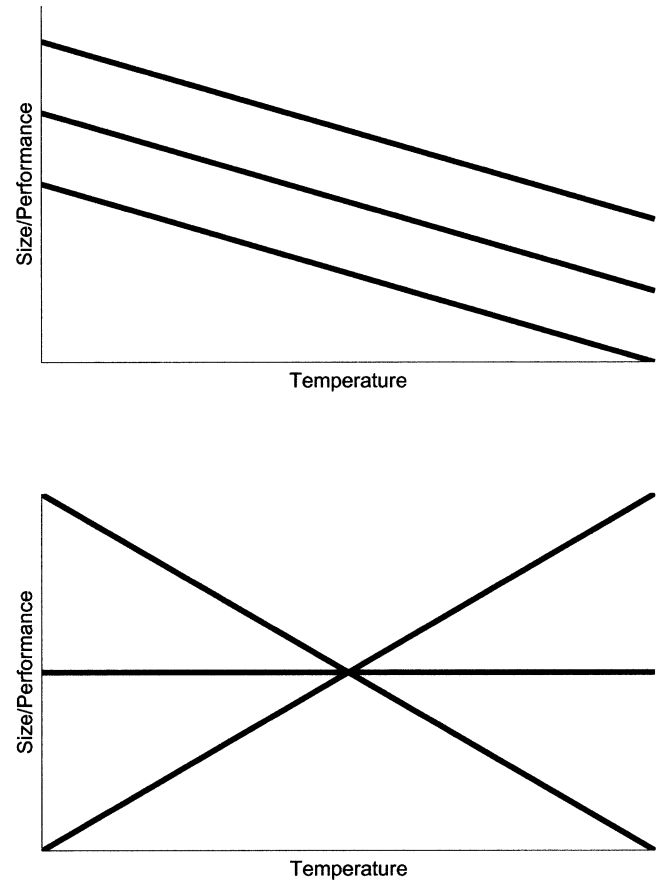


FIG. 1. Patterns of variation for linear thermal reaction norms or performance curves. *Upper panel:* Vertical shift, in which there is genetic variation in overall size or performance at all temperatures. *Lower panel:* variation in plasticity, in which there is genetic variation in the slopes of the reaction norms or performance curves.

etc) for the family or clone as a function of developmental temperature is the estimated DRN. In practice, the vast majority of studies of variation in thermal DRNs consider only 2 or 3 temperature conditions (but see David *et al.* [1994, 1997]; Karan *et al.* [1999]).

For a sample of clones, genetic lines or families from a population, we can partition the total genetic variation of thermal developmental reaction norms into two components: variation in the mean phenotypic value across all temperatures; and variation in thermal sensitivity (plasticity) of the phenotype at different temperatures. For simplicity consider the case of linear reaction norms or linear performance curves (or equivalently, where only two temperature conditions are measured). For example, the first component of variation represents genetic variation in overall size or performance independent of temperature—*i.e.*, genetic variation in the vertical position (Fig. 1, top). Because the reaction norms are parallel, there is no genetic variation in plasticity. The second component represents gene-environment interactions—*i.e.*, genetic variation in plasticity (Fig. 1, bottom) (Scheiner, 1993).

Because most performance curves (and many developmental reaction norms) are strongly non-linear,

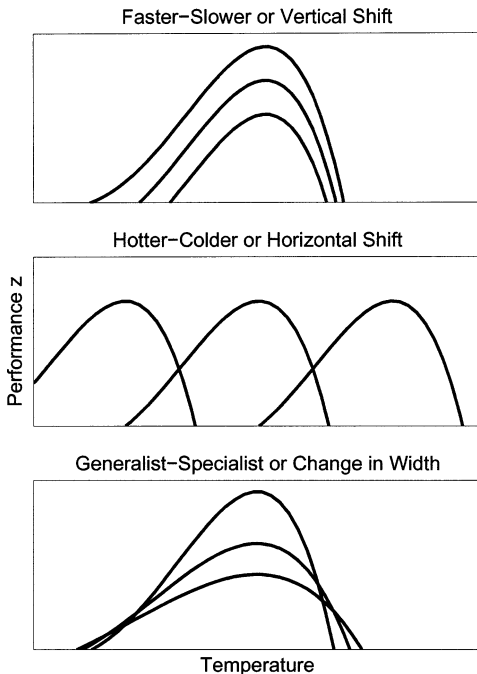


FIG. 2. Patterns of variation for nonlinear thermal performance curves. *Upper*: Vertical shift, in which there is genetic variation in performance at all temperatures. *Middle*: Horizontal shift, in which there is genetic variation in the position of the performance curve along the temperature axis. *Lower*: Generalist-Specialist, in which there is a genetic tradeoff between maximal performance and performance breadth.

the patterns of genetic architecture can be more complex. For example, evolutionary physiologists have proposed three interesting components of variation in thermal performance curves: vertical shift (faster-slower), horizontal shift (hotter-cooler) and generalist-specialist (Fig. 2) (Huey and Kingsolver, 1989). Vertical shift represents variation in overall performance across all temperatures; horizontal shift represents variation in the position of the performance curve along the temperature axis (*e.g.*, variation in the temperature at which performance is maximized); generalist-specialist represents the genetic tradeoff between maximal performance and performance breadth (Gilchrist, 1996; Huey and Kingsolver, 1989; Kingsolver *et al.*, 2001a; Palaima and Spitze, 2004). The vertical shift component represents genetic variation in overall performance or size independent of temperature, whereas the horizontal shift and generalist-specialist components represent two distinct types of gene-environment interactions—*i.e.*, genetic variation in plasticity.

These proposed patterns of variation in DRNs and TPCs (Fig. 1–2) can be viewed as the decomposition of variation in the data in predetermined directions or “modes” (Kingsolver *et al.*, 2001a). In particular, the vertical shift direction represents a vector with equal loadings at all temperatures; and decomposition of variation in this direction is similar to principal components analysis. The estimate of the total between-family variation or sums of squares,  $SST$ , is decomposed

into two parts,  $SS_v$  and  $SSE$ , such that  $SST = SS_v + SSE$ . Here  $SS_v$  estimates the variation along the vertical shift mode (it is the sums of squares of the variation in the projected data onto the vertical shift mode) and  $SSE$  is the residual sums of square. Then the proportion of variation along the vertical shift mode is then estimated by  $RSS_v = SS_v/SST$ . This provides a useful means to visualize and quantify genetic variation in overall size or performance, and to compare this to genetic variation in plasticity (see the Appendix for details).

What is the relative magnitude of genetic variation in overall size or performance (vertical shift) compared with genetic variation in plasticity? Does this differ for TDRNs and TPCs? To illustrate the approach, first consider two case studies. The studies of Scheiner and Lyman (Scheiner *et al.*, 1991; Scheiner and Lyman, 1989) provided a classic example of thermal developmental reaction norms of adult size in *Drosophila melanogaster*, and S. Scheiner has generously made their original data available to us. Their study used 47 half-sib families reared at two constant temperatures (19 and 25°C), and measured thorax length as an index of adult body size (Fig. 3, upper left). As in most ectotherms (Angilletta *et al.*, 2003; Atkinson, 1993), mean adult size decreased with increasing temperature, but there was substantial between-family (genetic) variation in DRNs. In our analyses, we removed the effect of the mean DRN (Fig. 3, upper right), and projected the data in the direction of the vertical shift mode, *i.e.*, the vector with equal loadings at all temperatures (Fig. 3, lower left). The projections clearly illustrate that there was substantial genetic (between-family) variation in overall size across temperatures. When the effects of these projections were removed, the residuals illustrate the variation in plasticity among families (Fig. 3, lower right). Least-squares estimates show that 69% of the total genetic variation represents genetic (between-family) variation in overall size (vertical shift), while the remaining 31% represents genetic variation in plasticity. Scheiner and Lyman used selection experiments to alter the mean plasticity in the population, confirming the existence of genetic variation in plasticity (Scheiner and Lyman, 1991).

Let us apply the same approach to analyzing thermal performance curves from a study of TPCs for short-term growth rate in *Pieris rapae* caterpillars (Kingsolver *et al.*, 2004). In this system, an individual caterpillar in the field may experience body temperatures ranging over 20–25°C in a single diurnal cycle, and thus express a wide range of feeding and growth rates over hourly to daily time scales (Kingsolver, 2000). In the study, the short-term (2–12 h) growth rate of each individual caterpillar was measured at 6 different temperatures between 11–35°C during a single larval instar (4th) in a full-sib design, with 10–12 individuals in each full-sib family (see Kingsolver *et al.*, 2004 for details). The resulting TPCs showed a strong effect of temperature on mean growth rate, with substantial variation among families (Fig. 4). Our projections illus-

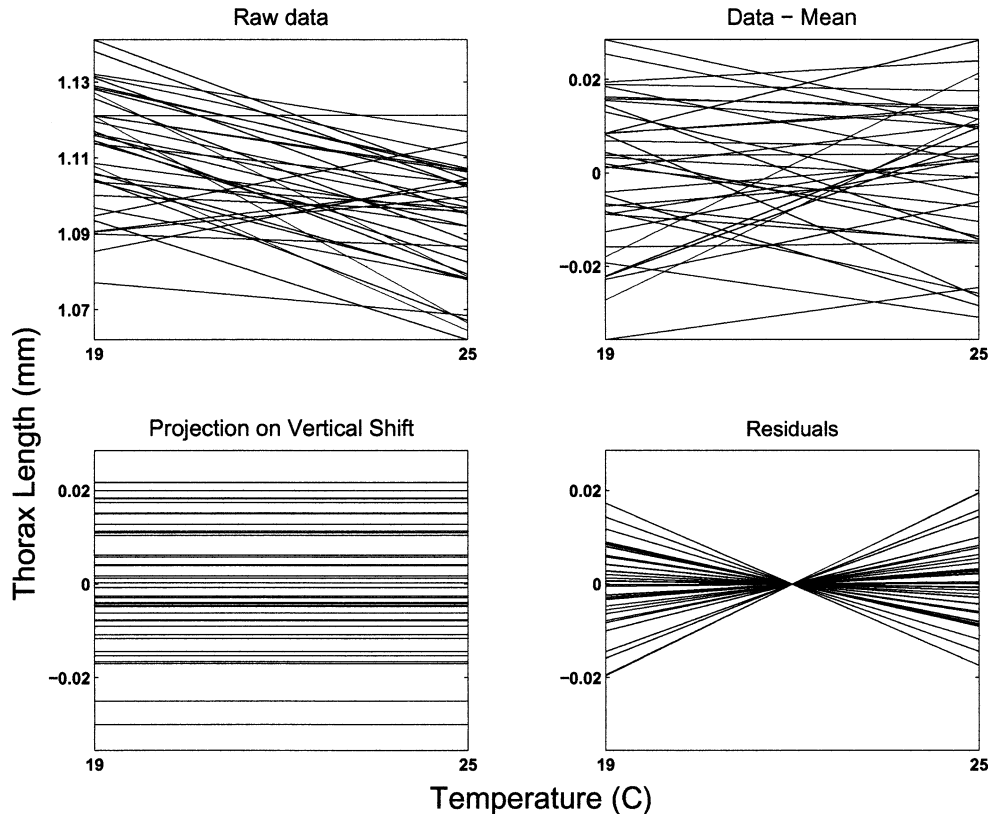


FIG. 3. Variation among sib-families of *Drosophila melanogaster* for developmental reaction norms (DRN) of adult thorax length at two developmental temperatures. Data from Scheiner and Lyman (1989). See text for explanation of the analyses. *Upper left*: Raw data for mean DRN of each family. *Upper right*: Data corrected for grand mean DRN. *Lower left*: projection of each family mean DRN in the vertical shift direction. *Lower right*: residuals after effects of vertical shift projections are removed, representing variation in plasticity.

trate that there was some variation in overall performance independent of temperature (Fig. 4, lower left), with substantial residual variation (Fig. 4, lower right). For these data, only 22% of the total genetic variation represents genetic variation in overall growth rate, and the remaining 78% represents genetic variation in plasticity.

Clearly the proportion of genetic variation associated with overall size and growth rate is quite different in these two examples: What is generating these different patterns of genetic architecture? Through the generosity of our colleagues, we were able to obtain a number of data sets on thermal developmental reaction norms and thermal performance curves for size and performance of insects. Each study included in our analysis considered genetic variation within a species for at least 20 families or genetic lines measured at two or more temperatures. There were 6 data sets for DRNs and 3 data sets for TPCs; the measured traits included aspects of adult body size, growth rate and locomotory speed. We estimated the % of total genetic (between-family or between-line) variation that represents variation in vertical shift (*i.e.*, variation in overall size or performance) in each study, using two different methods: the projection on the vertical shift mode described above; and a mixed-model ANOVA in which variance components were estimated via REML (Fry,

1992). The latter method makes additional assumptions about the data, but allows for the estimation of the error (environmental) variance and the within-family variation, in addition to the genetic components (see the Appendix).

The results of our analyses are summarized in Table 1. The studies varied substantially in the number of temperatures (from 2 to 7) and the range of mean temperatures (from 6°C to 31°C). One striking pattern is that estimated % of genetic variation associated with vertical shift was consistently higher for developmental reaction norms than for thermal performance curves—indeed, there was no overlap in the values. For the one system study (*Pieris rapae*) for which there were data for both TPCs and DRNs, the % of vertical shift was greater than 80% of the DRNs but less than 30% for the TPCs. More generally for DRNs, the % of vertical shift exceeded 50% in most cases, suggesting that variation in plasticity (gene-environment interactions) is typically less than half of the total genetic variation in for most DRNs. By contrast, for the three datasets for TPCs, the % of vertical shift was below 33% in most cases, suggesting that variation in plasticity comprises most of the total genetic variations for TPCs. The two methods gave qualitatively similar patterns, although the estimated % vertical shift differ considerably in a few cases (see Appendix). Note also



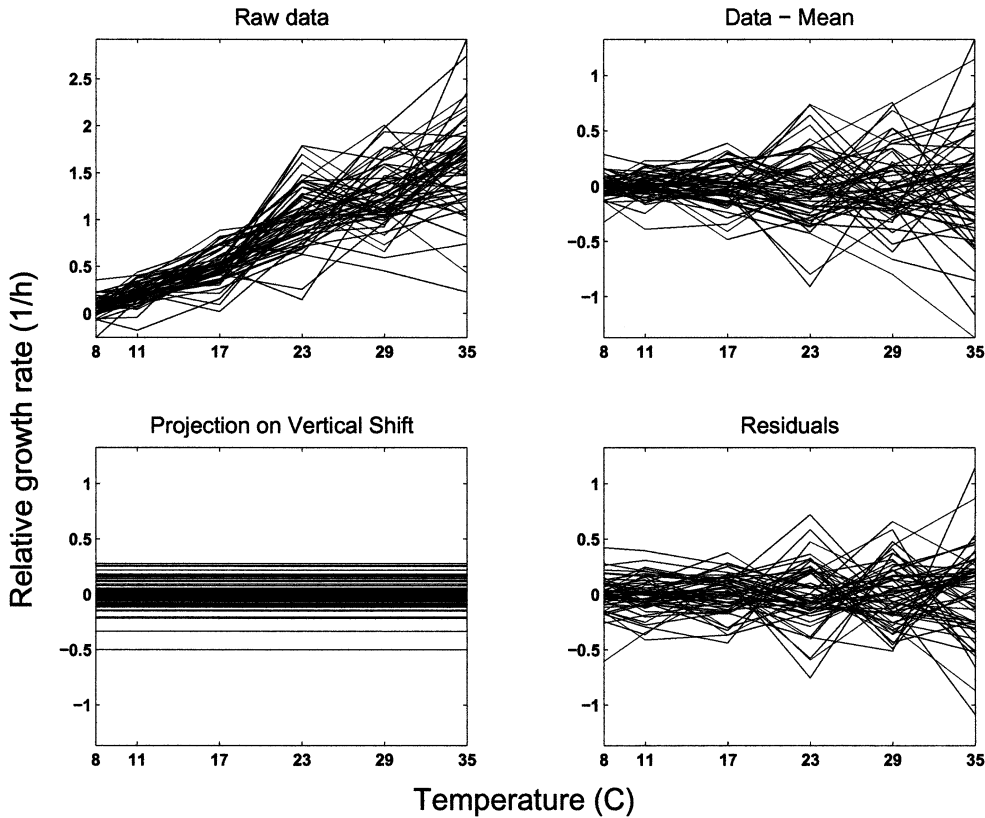


FIG. 4. Variation among full-sib families of *Pieris rapae* caterpillars for thermal performance curves (TPC) of short-term growth rate at 6 temperatures during the 4<sup>th</sup> larval instar. Data from Kingsolver *et al.*, 2004). See text for explanation of the analyses. *Upper left*: Raw data for mean TPC of each family. *Upper right*: Data corrected for grand mean TPC. *Lower left*: projection of each family mean TPC in the vertical shift direction. *Lower right*: residuals after effects of vertical shift projections are removed, representing variation in plasticity.

that the % of error variation from the REML analyses is considerably larger for TPCs than for DRNs, no doubt reflecting the greater behavioral variability and measurement error associated with shorter-term measures of performance compared with size and time measurements.

We can suggest four factors that probably contribute to these patterns. One factor is the number and range of temperatures measured. To illustrate this, we used one dataset for TPCs of growth rate in *P. rapae* caterpillars that considered 6 temperatures ranging from 8 to 35°C (see Table 1), and sampled subsets of these data. For the full dataset (6 temperatures) the estimated % of vertical shift was 21%; however choosing 2 or 3 temperatures increased the % of vertical shift to greater than 46%. When the temperatures considered are closely spaced (*e.g.*, 8 and 11°C) the % of vertical shift increased to 71%; this will occur if trait values at nearby temperatures are positively correlated genetically, as one would expect. Because many studies of DRNs have considered only 2 temperature environments over a relatively narrow range of mean temperatures, these may overestimate the genetic contributions of vertical shift and thereby underestimate the genetic variation in plasticity. Second, in the vertical projections (Appendix, eqn 9) method, measurement error may be confounded with estimates of gene-en-

vironment interaction, reducing the estimated % of vertical shift. As indicated in Table 1 such error is greater for measurements of performance than for morphometric traits. However, our REML estimates from the ANOVA analyses provide independent estimates of the genetic and error variance components, minimizing this problem. In fact, the REML estimates for % of vertical shift are even lower for two of the three TPC studies.

Third, the strongly non-linear nature of TPCs will contribute to the importance of gene-environment interactions and genetic variation in plasticity (Kingsolver *et al.*, 2001a). For example, horizontal shift variation in a non-linear performance curve (Fig. 2, middle) will generate strong gene-environment effects resulting in variation in plasticity. In contrast, horizontal shift in a perfectly linear reaction norm (Fig. 1) will cause no gene-environment effects because there is no variation in slopes: in this case variation in the x-intercept (horizontal shift) is directly equivalent to variation in the y-intercept (vertical shift). As a result, the degree of non-linearity (*e.g.*, the amount of total or mean curvature) in reaction norms and performance curves will tend to result in greater genetic variation in plasticity. Because thermal performance curves are frequently more strongly non-linear than developmen-

TABLE 1. Analyses of the % of vertical shift for development reaction norms (DRN) and thermal performance curves (TPC) of size and performance in insects.\*

Type	Species	Trait	% vertical (ANOVA)	% vertical (ANOVA)	% error (ANOVA)	Temperatures	# temps	Range	# of families	Source
DRN	<i>Blattella germanica</i>	growth (21 d)	77.16	72.34	29.43	20,20/30,30	3	10	41	1
DRN	<i>Drosophila melanogaster</i>	thorax size	68.75	52.57	86.71	19,25	2	6	37	2
DRN	<i>Drosophila melanogaster</i>	thorax length	56.96	NA	NA	12,14,17,21,25,28,31	7	19	20	3
DRN	<i>Drosophila melanogaster</i>	wing length	72.24	NA	NA	12,14,17,21,25,28,31	7	19	20	3
DRN	<i>Pieris rapae</i>	adult mass	86.63	92.70	55.78	10/34, 16/40	2	6	37	4
DRN	<i>Scathophaga stercoraria</i>	hind tibia length	61.01	75.27	55.86	12,18,24	3	12	31	5
DRN	<i>Scathophaga stercoraria</i>	hind tibia length	42.57	58.91	62.17	15,23	2	8	26	6
DRN	<i>Trimerotropis pallidipennis</i>	femur length	72.99	54.87	54.43	40/30, 30/25	2	7.5	87	7
TPC	<i>Aphidius ervi</i>	velocity	33.73	52.61	85.80	12,16,20,24,28,32,36	7	24	87	8
TPC	<i>Pieris rapae</i>	growth rate	21.71	0.00	83.19	8,11,17,23,29,35	6	27	56	9
TPC	<i>Pieris rapae</i>	growth rate	7.53	0.00	89.71	11,17,23,29,35,40	6	29	32	9

% vertical: from vertical mode analysis, eqn (10) in the Appendix.

% vertical (ANOVA): from REML mixed effects model, eqn (12).

% error (ANOVA): error variance as a % of total variance, from REML mixed effects model.

Temperatures: temperature treatments used (in C); fluctuating temperatures indicated by /.

# temps: number of temperature treatments used in the study.

Range: range of mean temperatures used in the temperature treatments.

# of families: mean # of families per treatment.

Source of data:

1. (M. Angilletta, unpubl.); 2. (Scheiner and Lyman 1989); 3. (Karan et al. 2000); 4. (J. G. Kingsolver, unpubl.); 5. (Hosken et al. 2000); 6. (W. Blanckenhorn, unpubl.); 7. (G. Davidowitz, unpubl.); 8. (Gilchrist 1996); 9. (Kingsolver et al. 2004).

\* See Appendix for details of the analyses, and legend below for explanation of key data fields.

tal reaction norms, TPCs may typically reveal greater variation in plasticity than DRNs.

Fourth, for a performance curve, performance is typically measured over short time scales at each temperature, such that performance for an individual may be measured for a set of different temperatures. By contrast, for a developmental reaction norm the trait value is typically defined in terms of the (constant or fluctuating) thermal conditions over some time period of development (see Introduction). One consequence of this difference in time scale is that performance may be possible over a wider range of temperatures than development. For example, short-term (2–24 h) growth rate for *P. rapae* caterpillars, may be readily measured over temperatures from 8 to 40°C, but survival and development to adulthood does not occur at constant temperatures below 11°C or above 31°C (Chen and Su, 1982; Gilbert, 1984; Kingsolver, 2000). In addition, developmental reaction norms are sometimes measured for diurnally fluctuating temperatures because these are more relevant to natural thermal conditions. In our analysis we have characterized these fluctuating conditions in terms of mean temperatures experienced during development (see Table 1). However, as the range of diurnal temperature fluctuations increases, this will increase the magnitude of positive genetic covariances in trait values (e.g., size or development times) across different mean temperatures, and thus increase the % of vertical shift estimated. Development times for insects often differ for fluctuating and equivalent mean constant temperatures (Worner, 1992), but the consequences of this for variation in developmental reaction norms has been not widely explored. The studies of David and colleagues with *Drosophila* provide a notable exception (David et al., 1994; Petavy et al., 2001a, b).

The relative contribution of these different factors is unknown. However, our analyses do suggest that the architecture of genetic variation and plasticity may differ for performance curves and developmental reaction norms in some interesting ways. It would be very useful to explore these differences more systematically in study systems where both of these types of phenotypic plasticity can be characterized. In addition, our analyses only distinguish vertical shift from genetic variation in plasticity, without distinguishing between different forms of plasticity such as horizontal shifts and generalist-specialist tradeoffs (Fig. 2). Addressing this issue requires new, non-linear statistical methods for partitioning the genetic variation for non-linear performance curves (Izem et al., 2003).

#### SELECTION ON SIZE AND DEVELOPMENT TIME IN NATURE

Because size affects many aspects of the biology and fitness of organisms, we might expect selection on size to be commonplace (Peters, 1983). Indeed, numerous studies have documented natural and sexual selection on aspects of size in natural populations (Blanckenhorn, 2000). But is selection on size differ-

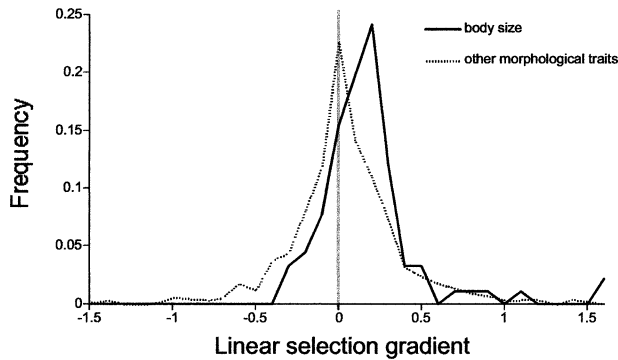


FIG. 5. Frequency distribution of linear selection gradients for phenotypic traits relating to overall size (solid line) and to other morphological traits (dotted line). From Kingsolver and Pfennig (2004).

ent in pattern or intensity than selection on other phenotypic traits? Similarly, the time to adulthood or to reproductive maturity is a key fitness component of an organism's life history (Stearns, 1992). Is there consistent selection on development time in nature?

Kingsolver and Pfennig (Kingsolver and Pfennig, 2004) have recently analysed field studies of selection to address these questions. The analyses used a database summarizing 14 years of published studies that estimated the strength of phenotypic selection on quantitative traits in natural populations in the field (Kingsolver *et al.*, 2001b). These studies measured the strength of directional selection in terms of the directional selection gradients ( $\beta$ ) (Lande and Arnold, 1983). The selection gradient  $\beta$  relates variation in the trait (in units of standard deviation of the trait) to variation in relative fitness (where mean fitness in a population or sample is defined to equal 1). The selection gradient thus provides a standardized metric of directional selection that facilitates comparisons among different traits and study systems.

Previous analyses of this database demonstrated that the frequency distribution of  $\beta$  was approximately symmetric about zero: *i.e.*, positive directional selection ( $\beta > 0$ ) was similar in frequency and magnitude to negative directional selection ( $\beta < 0$ ) (see *e.g.*, Fig. 1 in Kingsolver *et al.*, 2001b). This is not surprising, since we would not expect a consistent bias in the direction of selection for arbitrary quantitative traits (Kingsolver *et al.*, 2001b).

What about selection on overall size? Kingsolver and Pfennig (2004) subdivided the estimates of  $\beta$  for morphological traits into two classes: Size traits reflecting overall size (*e.g.*, body length, plant height, total mass, etc.); and Other Morphological traits. They identified 91 estimates of  $\beta$  for traits reflecting overall size, and 763 estimates of  $\beta$  for other morphological traits.

Their analyses of the frequency distributions of  $\beta$  for size and for other morphological traits showed a striking pattern (Fig. 5). For other morphological traits the distribution of  $\beta$  was approximately symmetric about zero, with 50% of the values greater than 0 and

a median value of 0.028. In contrast, the distribution of  $\beta$  for overall size was strongly shifted to positive values: 79% of the values were greater than 0 with the median  $\beta$  of 0.16. This same pattern holds when analyses are restricted to studies where selection on both size and other morphological traits is estimated with the same study. Interestingly, this qualitative pattern also held for different components of fitness (*e.g.*, survival, fecundity, or mating success) and for different taxa (invertebrates, plants and vertebrates) (Kingsolver and Pfennig, 2004).

What about patterns of selection on development time in nature? Using the database available for life history and phenological traits (Kingsolver *et al.*, 2001b), Kingsolver and Pfennig (2004) divided estimates into two categories: Development Time, for traits associated with age or time to adulthood, maturity, or reproduction; and Other life history (LH) traits, for the remaining life history and phenological traits. There were many fewer values than for size and other morphological traits, with 32 estimates of  $\beta$  for Development Time and 96 for Other LH traits. However these limited data suggested a clear pattern: estimates for Other LH traits were symmetric about zero (median = 0.01); but estimates for Development Time were strongly and significantly shifted towards negative values (median =  $-0.145$ ), with 84% of the estimates less than 0 (Kingsolver and Pfennig, 2004). We note however that most of these estimates for life history and development time are from plants, and the database included only a single field study of selection on development time for an ectothermic animal (Koenig and Albano, 1987). Clearly additional field studies of selection on development time in ectotherms are needed.

It is important to recognize that many excellent field studies of selection on size, development time and other traits do not report selection in terms of selection gradients, and therefore were not represented in Kingsolver and Pfennig's (2004) analyses. Nevertheless, the results suggest a strong tendency towards positive directional selection on overall size, in contrast with other morphological traits. Bigger size is associated with higher fitness in most natural populations that have been studied—that is, bigger is generally fitter. This provides quantitative support for the analyses of Blankenhorn (2000), who identified relatively few studies that demonstrated negative fitness consequences of large body size. More limited data, primarily for plants, suggest that shorter development times are consistently associated with higher fitness in natural populations. Taken together, selection favoring increased size and shorter development times implies that there should be selection for more rapid growth and development rates. Unfortunately, there are only a handful of studies that have directly measured selection on growth rates or other aspects of performance in the field (Brodie and Ridenhour, 2003; Kingsolver and Gomulkiewicz, 2003; Kingsolver *et al.*, 2001b; Kingsolver and Huey, 2003; Wikelski and Romero, 2003).

## SELECTION ON GROWTH AND SIZE IN FLUCTUATING ENVIRONMENTS

Adult size in insects and other ectotherms is largely the result of the larval (nymphal) growth rates integrated over the time period of development. If there is directional selection on adult size and/or development time, as shown in the previous section, this implies that there must be selection on aspects of growth rate and/or on the duration of the growth period. How is selection on larval growth rate related to selection on adult size? In particular in a fluctuating thermal environment, how does selection on thermal performance curves for larval growth rate affect selection on adult body size? To address this question we need to understand quantitatively both the thermal environment of individuals and the nature of selection on thermal performance curves.

As an example, consider the thermal environment of *Pieris rapae* caterpillars. Like most insect larvae and nymphs, *P. rapae* caterpillars do not thermoregulate behaviorally or physiologically (except to avoid exposure to deleteriously high body temperatures); as a result one can use physical models to quantify the thermal environment of caterpillars in the field (Bakken *et al.*, 1985; Kingsolver, 2000). Field measurements during the summer in Seattle WA USA show that individual caterpillars routinely experience a wide temperature range on a daily basis, with temperatures varying by 20–25°C or more over a single diurnal cycle (Kingsolver, 2000). As a result, an individual can express a wide range of growth rates—a range of values of its thermal performance curve,  $z(T)$ —in the field during its larval life.

Suppose that the specific temporal pattern of temperature variation is less important than the relative frequency of different temperatures in determining growth and size. Then a useful way to characterize the thermal environment is to quantify the frequency distribution of body temperatures,  $f(T)$ , experienced by caterpillars in the field. For example, Figure 6 shows  $f(T)$  for two populations of *P. rapae* caterpillars during summer (July–August) conditions. These temperature distributions are strongly skewed or bimodal, with a higher “nighttime” peak at lower temperatures, and a less distinct “midday” peak (or long tail) at higher temperatures (see Kingsolver, 2000, for additional examples). Because in these populations larval development times in summer are on the order of 3–5 weeks, these distributions illustrate the wide range of body temperatures experienced by individual caterpillars within a single larval generation. In addition, the strong skew or bimodality in  $f(T)$  implies that the mean temperature is actually experienced rather infrequently. For example, the mean temperatures for  $f(T)$  in WA and NC are 20 and 26°C, respectively, whereas the modal temperatures are 15 and 23°C (Fig. 6). As a result, mean temperatures may be an uninformative, even misleading, measure of the thermal environment.

How do we quantify selection on thermal perfor-

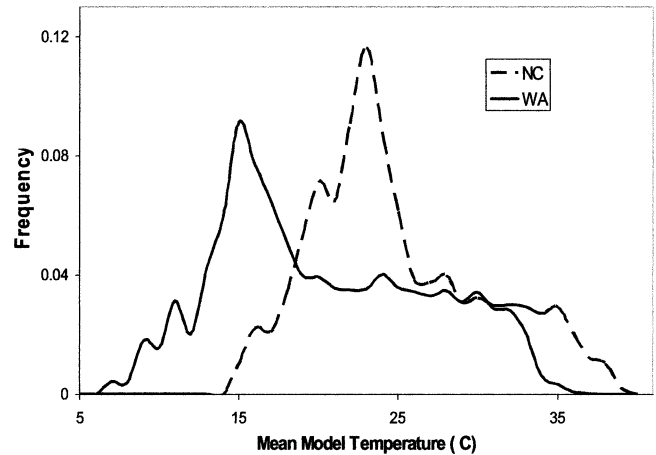


FIG. 6. Frequency distribution of mean operative temperatures of model *Pieris rapae* caterpillars in collard gardens, for July–August 2000 in Seattle WA (solid line) and July–August 2001 in Chapel Hill NC (dashed line). From Kingsolver and Gomulkiewicz (2003) and Kingsolver (unpublished).

mance curves? Kirkpatrick and colleagues extended the familiar multivariate model for evolution of quantitative traits to the case where the trait is a continuous function (Gomulkiewicz and Kirkpatrick, 1992; Kirkpatrick and Heckman, 1989). In this “function-valued” framework, selection is now represented by a selection gradient *function*,  $\beta(T)$ , that indicates the strength of directional selection on the performance curve  $z(T)$  over the continuous range of environments  $T$  (Beder and Gomulkiewicz, 1998; Gomulkiewicz, 2001; Gomulkiewicz and Kirkpatrick, 1992). If the strength of directional selection on performance changes with temperature  $T$ , then the selection gradient will be a function of temperature. As for other quantitative traits, the selection gradient for a TPC may be estimated by relating variation in the phenotype (*e.g.*, by measuring performance at a series of temperatures for each individual in the lab or field) to variation in relative fitness in the field (Kingsolver and Gomulkiewicz, 2003; Kingsolver *et al.*, 2001a).

Armed with quantitative measures of the thermal environment ( $f[T]$ ) and selection on a TPC for larval growth rate ( $\beta[T]$ ), we can now return to the relationship between selection on larval growth rate and selection on adult size in a fluctuating thermal environment. Kingsolver and Gomulkiewicz (2003) developed a simple theoretical model that explores this issue. The model assumes that the absolute fitness of an individual is directly proportional to its relative growth rate integrated over some relevant time period  $L$  (*e.g.*, over the duration of larval development or a selection episode). This assumption implies that relative growth rates at particular temperatures are not important except as they contribute to size at the end of the time period: *i.e.*, selection is acting on final size (or equivalently, on the development time required to achieve a fixed final size). Suppose further that growth rate and its contribution to fitness are determined only by tem-



perature, and not by the specific times at which that temperature occurs. This assumption implies that the frequency distribution  $f(T)$  (technically, the probability density function), rather than the temporal pattern of temperatures, is what determines growth rate and selection. With these assumptions there is a very simple relationship between the thermal environment and selection on a TPC (Kingsolver and Gomulkiewicz, 2003):

$$\beta(T) = Lf(T) \quad (1)$$

This simple result predicts that selection on the thermal performance curve is directly related to the frequency distribution of the thermal environment during selection. If selection acts directly on final (adult) body size, then the strength of directional selection on growth rate at different temperatures depends entirely on the relative frequency of those temperatures. For naturally fluctuating environments (Fig. 6), selection on growth rate will be strongest at the modal (most frequent) temperature, not at the mean temperature or temperatures at which growth rates are greatest (Kingsolver, 2000). These same predictions also apply to selection on development time and thermal performance curves for larval development rates.

Of course, one can imagine situations in which growth rate or performance at specific temperatures may contribute disproportionately to fitness. For example, if selection acts primarily on maximal growth rate or performance, then growth or performance at higher temperatures will contribute substantially to fitness even if such temperatures occur infrequently (Hertz *et al.*, 1988). We can relax the first assumption in the model above to consider this case, with the following result (Kingsolver and Gomulkiewicz, 2003):

$$\beta(T) = Lc(T)f(T) \quad (2)$$

where  $c(T)$  represents a weighting function that reflects the contributions of growth rate at temperature  $T$  to fitness. In this more general case, the strength of directional selection on thermal performance curves for growth rate depends on two factors: the frequency distribution of temperatures experienced during selection, and the weighting function relating growth at particular temperatures to fitness. Comparing eqns 1 and 2, we see that when selection acts only on total growth (final adult size), the weighting function  $c(T) = 1$  for all temperatures  $T$ .

As described by Kingsolver and Gomulkiewicz (2003), there are several interesting consequences of these theoretical predictions. First, selection on adult size or development time in fluctuating thermal environments can generate complex patterns of selection on thermal performance curves for growth or development rates, because of the bumpiness of  $f(T)$  in terrestrial environments of ectotherms (Fig. 6). Second, by comparing the selection gradient function ( $\beta[T]$ ) with the distribution of environmental conditions ( $f[T]$ ) in the field for some natural population, we can explore whether selection acts primarily on adult size

or development time, or on more complex aspects of thermal sensitivity of growth and development. For example, the failure of equation (1) to describe observed data may indicate that selection does not act simply on adult size, but instead acts disproportionately on growth rate at particular temperatures.

#### COMPARING REACTION NORMS AND PERFORMANCE CURVES

We have explored two aspects of plasticity for insects in response to temperature variation: thermal performance curves (TPCs) for growth rate, and thermal developmental reaction norms (TDRNs) for development time and adult body size. Our analyses suggest several interesting contrasting patterns between TPCs and TDRNs. First, for the datasets we considered, genetic variation in plasticity was relatively greater for performance curves than for developmental reaction norms. We suggest a number of possible methodological and biological reasons for this difference, but if this result is robust it has important implications for the evolution of plasticity for TPCs and TDRNs. Second, analyses of field selection studies indicate that there is consistent directional selection favoring increasing body size in most study systems. More limited data also indicate consistent directional selection favoring decreasing time to adulthood and reproductive maturity. These analyses suggest that in most systems studied thus far, larger size and faster growth rates are selectively favored. Third, adult body size is obviously a consequence of growth rates achieved throughout growth and development to the adult stage. We describe a theoretical model relating selection on adult size or development time to selection on thermal performance curves for growth rate. A key prediction of the model is that selection on growth rate is greatest at the most frequent temperatures, and not at mean temperatures or temperatures at which growth rate is maximized. These results illustrate the complex interplay between thermal performance curves of growth rate and thermal reaction norms of adult body size.

#### ACKNOWLEDGMENTS

We thank Mike Angilletta and Mike Sears for organizing the symposium. We thank Mike Angilletta, Wolf Blanckenhorn, Goggy Davidowitz, Jean David, Allen Gibbs, George Gilchrist, Brigitte Moreteau, and Sam Scheiner for their generosity in sharing published or unpublished data. The Whiteley Center at Friday Harbor provided the ideal place for analysis and writing. Research supported by NSF grants to JGK.

#### REFERENCES

- Angilletta, M. J., R. S. Wilson, C. A. Navas, and R. S. James. 2003. Tradeoffs and the evolution of thermal reaction norms. *Trends Ecol. Evol.* 18:234–240.
- Atkinson, D. 1993. Temperature and organism size—a biological law for ectotherms? *Adv. Ecol. Res.* 3:1–58.
- Atkinson, D. and R. M. Sibly. 1997. Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends Ecol. Evol.* 12:235–239.

- Bakken, G. S., W. R. Santer, and D. J. Erskine. 1985. Operative and standard operative temperature: Tools for thermal energetics studies. *Amer. Zool.* 25:933–943.
- Beder, J. H. and R. Gomulkiewicz. 1998. Computing the selection gradient and evolutionary response of an infinite-dimensional trait. *J. Math. Biol.* 36:299–319.
- Blanckenhorn, W. U. 2000. The evolution of body size: What keeps organisms small? *Quart. Rev. Biol.* 75:385–407.
- Brodie, E. D. and B. J. Ridenhour. 2003. Reciprocal selection at the phenotypic interface of coevolution. *Integr. Comp. Biol.* 43:408–418.
- Chen, C. N. and W. Y. Su. 1982. Influence of temperature on development and leaf consumption of three caterpillars on cauliflower *Artogeia rapae crucivora*, *Trichoplusia ni*, *Spodoptera litura*. *Chung Hua Chih Wu Pao Hu Husueh Hui Plant Prot. Bull.* 24:131–141.
- David, J. R., P. Gibert, E. Gravot, G. Petavy, J. P. Morin, D. Karan, and B. Moreteau. 1997. Phenotypic plasticity and developmental temperature in *Drosophila*: Analysis and significance of reaction norms of morphometrical traits. *J. Thermal Biol.* 22:441–451.
- David, J. R., B. Moreteau, J. P. Gauthier, G. Petavy, A. Stockel, and A. G. Imasheva. 1994. Reaction norms of size characters in relation to growth temperature in *Drosophila melanogaster*: An isofemale lines analysis. *Genetics, Selection, and Evol.* 26:229–251.
- de Jong, G. and P. Bijma. 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. *Livestock Product. Sci.* 78:195–214.
- Fry, J. D. 1992. The mixed-model analysis of variance applied to quantitative genetics: Biological meaning of the parameters. *Evolution* 46:540–550.
- Gilbert, N. 1984. Control of fecundity in *Pieris rapae* II. Differential effects of temperature. *J. Anim. Ecol.* 53:589–597.
- Gilchrist, G. W. 1996. A quantitative genetic analysis of thermal sensitivity in the locomotor performance curve of *Aphidius ervi*. *Evolution* 50:1560–1572.
- Gomulkiewicz, R. 2001. Selection and constraint in a continuously varying environment. In P. H. van Tienderen and P. M. Brakefield (eds.), *The evolution of phenotypic plasticity: The study of adaptive responses in a variable world*. Oxford University Press, Oxford. (In press)
- Gomulkiewicz, R. and M. Kirkpatrick. 1992. Quantitative genetics and the evolution of reaction norms. *Evolution* 46:390–411.
- Hertz, P. E., R. B. Huey, and T. Garland, Jr. 1988. Time budgets, thermoregulation, and maximal locomotor performance: are reptiles Olympians or Boy Scouts? *Amer. Zool.* 28:927–938.
- Hosken, D. J., W. U. Blanckenhorn, and P. I. Ward. 2000. Developmental stability in yellow dung flies (*Scathophaga stercoraria*): Fluctuating asymmetry, heterozygosity and environmental stress. *J. Evol. Biol.* 13:919–926.
- Huey, R. B. and D. Berrigan. 1996. Testing evolutionary hypotheses of acclimation. In I. A. J. a. A. F. Bennett (ed.), *Phenotypic and evolutionary adaptation to temperature*, pp. 205–237. Cambridge University Press, Cambridge, UK.
- Huey, R. B., D. Berrigan, G. W. Gilchrist, and J. C. Herron. 1999. Testing the adaptive significance of acclimation: A strong inference approach. *Amer. Zool.* 39:323–336.
- Huey, R. B. and J. G. Kingsolver. 1989. Evolution of thermal sensitivity of ectotherm performance. *Trends Ecol. Evol.* 4:131–135.
- Huey, R. B. and R. D. Stevenson. 1979. Integrating thermal physiology and ecology of ectotherms: A discussion of approaches. *Amer. Zool.* 19:357–366.
- Izem, R., J. S. Marron, and J. G. Kingsolver. 2003. Analyzing variation in thermal performance curves. *Integr. Comp. Biol.* 43:929.
- Karan, D., J. P. Morin, P. Gibert, B. Moreteau, S. M. Scheiner, and J. R. David. 2000. The genetics of phenotypic plasticity. IX. Genetic architecture, temperature, and sex differences in *Drosophila melanogaster*. *Evolution* 54:1035–1040.
- Karan, D., J. P. Morin, E. Gravot, B. Moreteau, and J. R. David. 1999. Body size reaction norms in *Drosophila melanogaster*: temporal stability and genetic architecture in a natural population. *Genetics, Selection, and Evol.* 31:491–508.
- Kingsolver, J. G. 2000. Feeding, growth and the thermal environment of Cabbage White caterpillars, *Pieris rapae* L. *Physiol. Biochem. Zool.* 73:621–628.
- Kingsolver, J. G. and R. Gomulkiewicz. 2003. Environmental variation and selection on performance curves. *Integr. Comp. Biol.* 43:470–477.
- Kingsolver, J. G., R. Gomulkiewicz, and P. A. Carter. 2001a. Variation, selection and evolution of function-valued traits. *Genetica* 112/113:87–104.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. H. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001b. The strength of phenotypic selection in natural populations. *The American Naturalist* 157:245–261.
- Kingsolver, J. G. and R. B. Huey. 2003. Introduction: The evolution of morphology, performance, and fitness. *Integr. Comp. Biol.* 43:261–366.
- Kingsolver, J. G. and D. W. Pfennig. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution* 58:1608–1612.
- Kingsolver, J. G., G. J. Ragland, and J. G. Shlichta. 2004. Quantitative genetics of continuous reaction norms: Thermal sensitivity of caterpillar growth rates. *Evolution* 58:1521–1529.
- Kirkpatrick, M., and N. Heckman. 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* 27:429–450.
- Koenig, W. D. and S. S. Albano. 1987. Lifetime reproductive success, selection, and the opportunity for selection in the white-tailed skimmer *Plathemis lydia* (Odonata: Libellulidae). *Evolution* 41:22–36.
- Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Palaima, A. and K. Spitze. 2004. Is a jack-of-all-temperatures a master of none? An experimental test with *Daphnia pulicaria*. *Evol. Ecol. Res.* 6:215–225.
- Petavy, G., J. R. David, P. Gibert, and B. Moreteau. 2001a. Viability and rate of development at different temperatures in *Drosophila*: a comparison of constant and alternating regimes. *J. Thermal Biol.* 26:29–39.
- Petavy, G., B. Moreteau, P. Gibert, J. P. Morin, and J. R. David. 2001b. Phenotypic plasticity of body size in *Drosophila*: Effects of a daily periodicity of growth temperature in two sibling species. *Physiol. Entomol.* 26:351–361.
- Peters, R. H. 1983. *The ecological implications of body size*. Cambridge University Press, Cambridge, UK.
- Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. *Annu. Rev. Ecol. Syst.* 24:25–68.
- Scheiner, S. M., R. L. Caplan, and R. F. Lyman. 1991. The genetics of phenotypic plasticity III. Genetic correlations and fluctuating asymmetries. *J. Evol. Biol.* 4:51–68.
- Scheiner, S. M. and R. F. Lyman. 1989. The genetics of phenotypic plasticity I. Heritability. *J. Evol. Biol.* 2:95–107.
- Scheiner, S. M. and R. F. Lyman. 1991. The genetics of phenotypic plasticity II. Response to selection. *J. Evol. Biol.* 4:23–50.
- Schlichting, C. D. and M. Pigliucci. 1998. *Phenotypic evolution: A reaction norm perspective*. Sinauer Associates, Sunderland, Massachusetts.
- Stearns, S. C. 1989. The evolutionary significance of phenotypic plasticity. *BioScience* 39:436–445.
- Stearns, S. C. 1992. *The evolution of life histories*. Oxford University Press, Oxford, UK.
- Via, S. and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39:505–522.
- Wikelski, M. and L. M. Romero. 2003. Body size, performance and fitness in Galapagos marine iguanas. *Integr. Comp. Biol.* 2003:376–386.
- Worner, S. P. 1992. Performance of phenological models under variable temperature regimes: consequences of the Kaufmann or rate summation effect. *Environ. Entomol.* 21:689–699.

## APPENDIX

Let  $z_{ijk}$  be the trait of interest (e.g., size or performance) of the  $k$ 'th individual in the  $i$ 'th family ( $1 \leq i \leq n$ , where  $n$  is the total

number of families) at temperature  $t_j$  ( $1 \leq j \leq d$ , where  $d$  is the total number of temperature measurements). We denote by  $K_{ij}$  the number of individuals in family  $i$  at temperature  $j$ , so ( $1 \leq k \leq K_{ij}$ ).

### 1 Projection on vertical shift

To graphically represent the raw data, we plotted the mean family trait vectors  $\mathbf{z}_i$ 's ( $\mathbf{z}_i = (z_{i1}, z_{i2}, \dots, z_{id})$ ) at temperatures ( $t_1, t_2, \dots, t_d$ ). Where  $z_{ij}$  is the mean trait across individuals in the same family  $i$  at temperature  $j$ , i.e.,

$$z_{ij} = \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} z_{ijk} \quad (\text{A1})$$

**Subtracting the mean.** The family vector mean is  $\bar{\mathbf{z}} = (z_{.1}, z_{.2}, \dots, z_{.d})$  such that

$$z_{.j} = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^{K_{ij}} \frac{1}{K_{ij}} z_{ijk} \quad (\text{A2})$$

We graphically represented the centered data by plotting the vectors  $\mathbf{z}_i - \bar{\mathbf{z}} = (z_{i1} - z_{.1}, z_{i2} - z_{.2}, \dots, z_{id} - z_{.d})$  at temperatures ( $t_1, t_2, \dots, t_d$ ).

**Projection on the vertical shift.** Let us denote by  $\mathbf{1}_d$  the  $d$ -vector of ones ( $1, 1, \dots, 1$ ). Then  $\bar{\mathbf{z}} + v_j \times \mathbf{1}_d$  represents a vertical shift variation of size  $v_j$  from the family mean  $\bar{\mathbf{z}}$ . So,  $\mathbf{1}_d$  is the vector representing the vertical shift. The projection matrix onto the vertical shift direction is the  $d \times d$  matrix of ones normalized by  $d$  to be of norm 1, i.e.,

$$J = \frac{1}{d} \mathbf{1}'_d \mathbf{1}_d = \frac{1}{d} \mathbf{1}_{(d,d)} \quad (\text{A3})$$

We can decompose the data into two orthogonal terms, the first represents the projection onto the vertical shift and the second represents the error, i.e.,

$$\mathbf{z}_i - \bar{\mathbf{z}} = \mathbf{proj}_i + \mathbf{e}_i$$

such that,

$$\mathbf{proj}_i = J \cdot (z_i - \bar{z})$$

and

$$\mathbf{e}_i = (I_{d,d} - J) \cdot (z_i - \bar{z}),$$

where  $I_{d,d}$  is the  $d \times d$  identity matrix.

We can simplify the projection term further as,

$$\begin{aligned} \mathbf{proj}_i &= \frac{1}{d} \sum_{j=1}^d (z_{ij} - z_{.j}) \mathbf{1}_d \\ &= (z_{i..} - z_{...}) \mathbf{1}_d \end{aligned}$$

**Quantifying the vertical shift by projections.** The variation along the vertical shift is quantified by the sums of squares  $SS_v$ , such that

$$SS_v = \sum_{i=1}^n \mathbf{proj}_i' \cdot \mathbf{proj}_i, \quad (\text{A4})$$

We have that

$$\mathbf{proj}_i' \cdot \mathbf{proj}_i = d(z_{i..} - z_{...})^2 \quad (\text{A5})$$

$$\mathbf{proj}_i' \cdot \mathbf{proj}_i = d(z_{i..} - z_{...})^2 \quad (\text{A6})$$

So,

$$SS_v = d \sum_{i=1}^n (z_{i..} - z_{...})^2 \quad (\text{A7})$$

$$SS_v = d \sum_{i=1}^n (z_{i..} - z_{...})^2 \quad (\text{A8})$$

On the other hand, the total variation SST is such that

$$\begin{aligned} SST &= \sum_{i=1}^n (\mathbf{z}_i - \bar{\mathbf{z}})' (\mathbf{z}_i - \bar{\mathbf{z}}) \\ &= \sum_{i=1}^n \sum_{j=1}^d (z_{ij} - z_{.j})^2 \end{aligned}$$

The percent of variation along the vertical shift is then quantified by the ratio  $RSS_v$  of the sums of squares  $SS_v$  on the total sums of squares SST, such that

$$RSS_v = \frac{SS_v}{SST} \quad (\text{A9})$$

$$= \frac{d \sum_{i=1}^n (z_{i..} - z_{...})^2}{\sum_{i=1}^n \sum_{j=1}^d (z_{ij} - z_{.j})^2} \quad (\text{A10})$$

The ratio  $RSS_v$  represents the **percent of vertical shift variation** estimated by projection on the vertical shift.

### 2 Analysis of variance models

We will see in this section that the  $RSS_v$  represents the fraction of the family effect (between family variation) on the family and family  $\times$  temperature effects in an ANOVA model.

**ANOVA model.** As in the previous section, let  $z_{ijk}$  be the value of the trait of interest of the  $k$ 'th individual in the  $i$ 'th family at temperature  $t_j$ . The model with two factors (family and temperature) with interaction (family  $\times$  temperature) is written as

$$z_{ijk} = \mu + f_i + t_j + (ft)_{ij} + \epsilon_{ijk} \quad (\text{A11})$$

where the overall mean is  $\mu$ , the family effect is  $f_i$ , the temperature effect is  $t_j$ , the interaction effect is  $(ft)_{ij}$  and the error term is  $\epsilon_{ijk}$ . Whether the effects are fixed or random, we have the following decomposition of the total sums of squares SST in the model

Source of variation	Sum of Squares
Due T	$SS_t$
Due F	$SS_f$
Due FT	$SS_{ft}$
Residual	SSE
Total	$SST = SS_t + SS_f + SS_{ft} + SSE$

The vertical shift ratio  $RSS_v$  that we defined in the previous section is equivalent to the fraction of the family effect on the family and family\*temperature effect, i.e.,  $RSS_v = SS_f / (SS_f + SS_{ft})$ .

**Quantifying the vertical shift in a mixed effect model.** For a fixed temperature effect, a random family effect and random interaction effect, the model defined by (11) is a mixed effect model. The mixed effect model assumptions we made in the analysis presented in this paper are

1. All  $\epsilon_{ijk}$ 's are independent from each other and they are sampled from a normal distribution of mean 0 and variance  $\sigma_\epsilon^2$ , i.e.,  $\forall 1 \leq i \leq n, \quad 1 \leq j \leq d, \quad 1 \leq k \leq K, \quad \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$
2. All  $f_i$ 's are independent from each other and they are sampled from a normal distribution of mean 0 and variance  $\sigma_f^2$ , i.e.,  $\forall 1 \leq i \leq n, \quad f_i \sim N(0, \sigma_f^2)$
3. All  $(ft)_{ij}$ 's are independent of each other and are sampled from a normal distribution of mean 0 and variance  $\sigma_{ft}^2$ , i.e.,  $\forall 1 \leq i \leq n, \quad 1 \leq j \leq d, \quad (ft)_{ij} \sim N(0, \sigma_{ft}^2)$
4.  $\epsilon_{ijk}$ 's,  $f_i$ 's, and  $(ft)_{ij}$ 's are all independent of each other.

The variances  $\sigma_f^2$ ,  $\sigma_{ft}^2$  and  $\sigma_\epsilon^2$  in the model are estimated by a Restricted Maximum Likelihood (REML) method. The **percent of vertical shift variation** using a mixed effect model is then estimated by the ratio

$$RVAR_v = \frac{\sigma_f^2}{\sigma_{ft}^2 + \sigma_f^2} \quad (\text{A12})$$

The ratio  $RVAR_v$  is, in general, different from  $RSS_v$ . In particular, a large within family variation (estimated by SSE) will cause a large difference between the two ratios.