doi:10.1111/evo.13342

The quantitative genetic basis of clinal divergence in phenotypic plasticity

Belinda van Heerwaarden[1,2] and Carla M. Sgrò[1]

[1]*School of Biological Sciences, Monash University Clayton 3800, Victoria, Australia*

[2]*E-mail: belinda.vanheerwaarden@monash.edu*

Received July 3, 2015
Accepted August 16, 2017

Phenotypic plasticity is thought to be an important mechanism for adapting to environmental heterogeneity. Nonetheless, the genetic basis of plasticity is still not well understood. In *Drosophila melanogaster* and *D. simulans*, body size and thermal stress resistance show clinal patterns along the east coast of Australia, and exhibit plastic responses to different developmental temperatures. The genetic basis of thermal plasticity, and whether the genetic effects underlying clinal variation in traits and their plasticity are similar, remains unknown. Here, we use line-cross analyses between a tropical and temperate population of *Drosophila melanogaster* and *D. simulans* developed at three constant temperatures (18°C, 25°C, and 29°C) to investigate the quantitative genetic basis of clinal divergence in mean thermal response (elevation) and plasticity (slope and curvature) for thermal stress and body size traits. Generally, the genetic effects underlying divergence in mean response and plasticity differed, suggesting that different genetic models may be required to understand the evolution of trait means and plasticity. Furthermore, our results suggest that nonadditive genetic effects, in particular epistasis, may commonly underlie plastic responses, indicating that current models that ignore epistasis may be insufficient to understand and predict evolutionary responses to environmental change.

KEY WORDS: **Body size, curvature, elevation, reaction norm, slope, thermal.**

Phenotypic plasticity, the ability of a genotype to alter its phenotype under different environmental conditions (Bradshaw 1965), is thought to be an important mechanism for responding to heterogeneous environments (Janzen 1967; Sultan and Spencer 2002; Ghalambor et al. 2004), and may be significant for population persistence in the face of novel environmental change (Chevin et al. 2013). Additionally, while phenotypic plasticity allows an organism to counter environmental changes, canalization, where a phenotype remains constant under different environments, may also be an important mechanism that buffers phenotypes against environmental perturbations (Waddington 1942). As such, phenotypic plasticity and canalization describe different facets of the same phenomenon: the sensitivity of phenotype to the environment (Debat and David 2001). Plastic responses are prevalent in nature, and the large number of studies reporting genotype-by-environmental (G × E) interactions within and/or between populations suggests that genetic variation for plasticity is widespread (Scheiner 1993; Kruuk et al. 2008; Des Marais et al. 2013). Nonetheless, the genetic basis, and the quantitative genetic architecture of plasticity/canalization is still not well understood, particularly in natural populations (Scheiner 1993; Via 1993; Via et al. 1995; Flatt 2005).

Two main models have been proposed to describe the genetic mechanisms underlying plastic responses (reviewed in Scheiner 1993; Via et al. 1995). The first model, the allelic sensitivity, or pleiotropic model, suggests that phenotypic plasticity is a function of differential expression of the same genes under different environments (Falconer 1952; Via and Lande 1985; Via 1993). Thus, in different environments similar loci will respond, but individual alleles may vary in their sensitivity/expression. The second model, the gene regulation or epistasis model, suggests that plasticity is due to genes that determine the magnitude of responses to environmental effects that interact with genes that determine the average expression of the character (Scheiner and Lyman 1989; Scheiner 1993; Schlichting and Pigliucci 1993). Although there has been some controversy over whether plasticity is underpinned by specific "plasticity genes" or through environmentally sensitive alleles (Via 1993; Scheiner 1993; Schlichting and Pigliucci 1993), mapping, expression, and family/selection studies have found empirical evidence for both models (e.g., allelic sensitivity:

Barnes et al. 1989; Weber and Scheiner 1992; Wu 1998; Ungerer et al. 2003; Lacaze et al. 2009; e.g., gene regulation: Li et al. 2010; Zhou et al. 2012; Mendez-Vigo et al. 2016). Furthermore, de Jong (1995) suggested that both models are mathematically equivalent, and may be modeled using a reaction norm approach (i.e., the expression of a character as a function of an environmental variable). Using this approach, de Jong and Gavrilets (2000) found that changes in the genetic variance and covariance of reaction norm parameters under different levels of environmental variance depend on the number of pleiotropic loci, suggesting that the gene regulation and allelic sensitivity models will have different consequences for the evolution of the genetic variance in reaction norm parameters (de Jong and Gavrilets 2000) and thus on the rate of evolution of plasticity.

The relative importance of additive and nonadditive genetic variation in adaptive evolution has been the subject of much controversy, with some researchers suggesting that the evolution of quantitative traits will occur predominately via additive gene action (Fisher 1930; Coyne et al. 1997, 2000; Hill et al. 2008), whereas others propose that nonadditive effects may be important (Wright 1931; Fenster et al. 1997; Wade and Goodnight 1998; Paixao and Barton 2016). The extent to which adaptive evolution occurs predominately via additive or nonadditive effects has direct implications for speciation models and the evolution of reproductive isolation (Dobzhansky 1937; Mayr 1954; Carson and Templeton 1984; Orr 1995), the evolution of sex and recombination (Maynard Smith 1978; Barton 1995), the maintenance of genetic variation (Gimelfarb 1989; Hermisson et al. 2003), animal and plant breeding (Lee and Kim 2009; Fethi et al. 2011), and conservation genetics (Fenster et al. 1997). Significantly, epistasis plays an important role in evolutionary models of canalization and genetic robustness (Wagner et al. 1997; Rice 1998; Flatt 2005). Furthermore, a study on *Escherichia coli* found that mutations that show epistasis were disproportionately likely to also show phenotypic plasticity (Remold and Lenski 2004), suggesting that epistasis may generally be fundamental to the evolution of environmental sensitivity (plasticity and canalization). Nonetheless, we currently know little about the quantitative genetic architecture of environmental sensitivity/plasticity.

Although the gene regulation/epistasis model of plasticity emphasizes physiological epistatic gene interactions between regulatory and trait loci, it is unknown whether these interactions generate epistatic variance and influence the evolution of plasticity (Via et al. 1995). Several quantitative trait locus (QTL) studies have found evidence for dominance and/or epistatic genetic effects for environmental sensitivity/plasticity (Wu 1998; Li et al. 2014; Mendez-Vigo et al. 2016), suggesting that nonadditive genetic effects may underpin plastic responses. However, the extent to which these loci generate epistatic variance (rather than additive variance) that influences the evolution of plasticity, especially in natural populations, is not clear. Although line-cross analyses (Mather and Jinks 1982) and crosses examining outbreeding depression (or F_2 breakdown) (Fenster et al. 1997) do not measure standing levels of additive, dominant, and epistatic genetic variance, and thus cannot predict the contribution of nonadditive genetic effects to any future short-term responses to selection, they can detect whether epistatic interactions were important in the evolution of divergence in means across populations/lines (Fenster et al. 1997). A small number of controlled crossing experiments have found evidence that dominance and/or epistatic genetic effects underlie divergence in environmental sensitivity/plasticity (Westerman 1971a, b; Perkins and Jinks 1973; Connolly and Jinks 1975; Pooni et al. 1987). However, most studies were conducted on inbred or artificially selected lines, rather than natural populations/populations recently collected from the field (but see Westerman 1971a). Furthermore, past studies have also failed to clearly distinguish contributions from dominance compared to different types of digenic epistatic effects; nor were they able to account for maternal effects, which if present, may lead to the erroneous detection of epistasis (Gilchrist and Partridge 1999; Kennington et al. 2001). Importantly, to our knowledge, no studies have used a comprehensive crossing design to investigate the genetic architecture underlying divergence in plasticity in populations collected from different biogeographical locations. This information would provide valuable insight into the extent to which nonadditive genetic effects are important for adaptive evolution of plasticity in nature, and whether current plasticity models that assume that additive genetic effects underlie evolution are sufficient.

For *Drosophila*, temperature is a major environmental factor influencing the geographic distribution of species, and adult tolerance to thermal extremes (heat and cold resistance) provides a good predictor of current species distributions (Overgaard et al. 2014). In *D. melanogaster*, opposing genetic clines in heat and cold resistance have been observed along the east coast of Australia; tropical populations have higher heat resistance than temperate populations, whereas temperate populations recover faster from a cold stress than tropical populations (Hoffmann et al. 2002; Sgro et al. 2010; Cockerell et al. 2014). Similar to many other species (Meiri and Dayan 2003; Chown and Gaston 2010), body size also varies with latitude, with size increasing at higher latitudes (James et al. 1995; van Heerwaarden and Sgro 2011).

In *Drosophila*, body size, heat, and cold resistance also respond plastically to temperature. Warmer rearing temperatures result in smaller flies (Atkinson 1994; James et al. 1997; Azevedo et al. 1998; van Heerwaarden and Sgro 2011) and increased heat resistance, whereas cooler rearing temperatures commonly increase cold resistance and size (Hoffmann et al. 2005; Cockerell et al. 2014). Although clinal patterns for the mean value of these

traits in two species, *D. melanogaster* and *D. simulans*, do not change significantly across different developmental temperatures, there is evidence for genetically based differences in plasticity for body size and heat and cold resistance along the eastern Australia cline in both species (van Heerwaarden and Sgro 2011; Cockerell et al. 2014). Furthermore, as phenotypically plastic responses to temperature mirror clinal patterns (i.e., plastic changes in these traits in response to rearing temperature change in the same direction as genetic patterns in response to temperature changes along the latitudinal gradient), these responses may be adaptive (Huey and Berrigan 1991; Hoffmann et al. 2005; Fallis et al. 2014). Line-cross and clinal analyses have been performed between tropical and temperate populations of *D. melanogaster* and *D. simulans* from the east coast of Australia to investigate whether clinal patterns and the genetic effects contributing to clinal divergence in mean wing centroid size, thorax length, wing-to-thorax ratio, cold and heat resistance differ under different developmental temperatures (18°C, 25°C, and 29°C) (van Heerwaarden and Sgro 2011). Specifically, we (van Heerwaarden and Sgro 2011) showed that the genetic basis of these traits is environment specific. However, we did not explore the genetic basis of divergence in the thermal plasticity of these traits.

The aim of this study was to use line-cross analyses to investigate the quantitative genetic basis of clinal divergence in plasticity for thermal stress resistance (heat and cold resistance) and body size (wing centroid size and thorax size) in *D. melanogaster* and *D. simulans*. Specifically, as current plasticity evolution models ignore nonadditive genetic variance, we were interested in determining the extent to which additive and nonadditive genetic effects, particularly epistasis, underlie divergence in plastic responses. We took a reaction norm approach to studying plasticity, where different functions (parameters) of the reaction norm can be used to analyze average performance (trait mean, elevation), and the degree of plasticity—(slope) and the shape (curvature)—of the reaction norm (David et al. 1997; Berger et al. 2014; Murren et al. 2014). This approach enabled us to investigate whether the genetic basis of divergence in mean performance of each trait (elevation) was similar to the genetic basis of divergence in plasticity (slope and curvature). Although line-cross analyses cannot directly distinguish which genetic model (allelic sensitivity or genetic regulation hypothesis) may underlie clinal divergence in plasticity, as the gene regulation genetic model involves physiological epistasis, we were interested in investigating whether statistical epistasis contributed more frequently to divergence in environmental sensitivity (plasticity/canalization) than mean performance. Finally, we also explored whether the genetic effects underlying curvature and slope differed to explore whether different aspects of plastic responses evolve via different genetic mechanisms to further understand the evolution of plasticity.

Methods

EXPERIMENTAL POPULATIONS

Populations of *D. melanogaster* and *D. simulans* were collected from a tropical (Gordenvale, north-eastern Queensland, 17°10′05″S, 145°49′55″E) and temperate (Melbourne, Victoria, 37°47′30″S, 145°26′05″E) location in January 2008. Twenty single field-collected females of each species from each location were used to find 20 isofemale lines for each population. Three generations after field collection, a mass-bred population was initiated with 20 males and 20 females from each of the 20 isofemale lines, per species per location. Each mass-bred population was maintained at 25°C under a 12-h light:12-h dark cycle at a census population size of approximately 1000 individuals across 3 × 250 mL bottles containing 20 mL of potato, yeast, and sucrose media. Line crosses were performed after six (*D. simulans*) or nine (*D. melanogaster*) generations of mass breeding.

EXPERIMENTAL DESIGN

Line crosses were performed to examine the relative contribution of different composite genetic effects (CGEs, e.g., additive, dominance, and epistatic gene effects) to the divergence in mean performance and plasticity for heat knockdown time, chill coma recovery time (*D. melanogaster*), wing centroid size, and thorax length between tropical and temperate populations of *D. melanogaster* and *D. simulans*. Line crosses are a widely used quantitative genetics method for estimating the genetic architecture underlying divergence in a phenotype of interest between two strains or populations. This approach involves crossing two parental strains to produce an F_1 cohort, and performing subsequent crosses (e.g., F_2, backcrosses, reciprocal crosses) to generate cohorts that have different combinations of parental genes. The observed mean phenotypic values of the parental and subsequent cross cohorts are then compared to the means expected from different genetic models that include additive, dominance, epistatic, maternal, and/or cytotype effects (maternal effects inherited from the organelles [e.g., mitochondria] or microorganisms in the cytoplasm) (Mather and Jinks 1982; Kearsey and Pooni 1996; Fox et al. 2004).

To generate cohorts that have different combinations of parental genes, we set up 14 crosses, similar to the procedure outlined in Gilchrist and Partridge (1999) (see Table 1), and as described in van Heerwaarden and Sgro (2011). Crosses were initiated between the two parental lines (temperate and tropical populations of each species) and the subsequent F_1 and F_2 generations. The F_1 generation was then backcrossed to the parents. Each cross included a reciprocal cross. Reestablishing the parental and F_1 crosses each generation allowed all 14 cohorts and parentals to be tested simultaneously after three generations of crossing. All crosses were performed at 25°C, and were initiated with 100 virgin females and 100 males. In the third generation, when all

Table 1. Outline of crossing scheme (performed separately for each species), the relative contribution of the CGEs to each cohort, and the proportion of P1 genes.

Cohort	Cross (dam × sire)	*m*	*Aa*	*Ad*	*AaAa*	*AaAd*	*AdAd*	*Mea*	*Med*	*Ca*	*CaAa*	*CaAd*	Proportion of P1 genes
P1	Temperate population	1	1	0	1	0	0	1	0	1	1	0	1
P2	Tropical population	1	−1	0	1	0	0	−1	0	−1	1	0	0
F_1	P1 × P2	1	0	1	0	0	1	1	0	1	0	1	0.5
F_1R	P2 × P1	1	0	1	0	0	1	−1	0	−1	0	−1	0.5
F_2	(P1 × P2) × (P1 × P2)	1	0	0.5	0	0	0.25	0	1	1	0	0.5	0.5
F_2R	(P2 × P1) × (P2 × P1)	1	0	0.5	0	0	0.25	0	1	−1	0	−0.5	0.5
B1a	P1 × (P1 × P2)	1	0.5	0.5	0.25	0.25	0.25	1	0	1	0.5	0.5	0.75
B1b	P1 × (P2 × P1)	1	0.5	0.5	0.25	0.25	0.25	1	0	1	0.5	0.5	0.75
B1Ra	(P1 × P2) × P1	1	0.5	0.5	0.25	0.25	0.25	0	1	1	0.5	0.5	0.75
B1Rb	(P2 × P1) × P1	1	0.5	0.5	0.25	0.25	0.25	0	1	−1	−0.5	−0.5	0.75
B2a	P2 × (P1 × P2)	1	−0.5	0.5	0.25	−0.25	0.25	−1	0	−1	0.5	−0.5	0.25
B2b	P2 × (P2 × P1)	1	−0.5	0.5	0.25	−0.25	0.25	−1	0	−1	0.5	−0.5	0.25
B2Ra	(P1 × P2) × P2	1	−0.5	0.5	0.25	−0.25	0.25	0	1	1	−0.5	0.5	0.25
B2Rb	(P2 × P1) × P2	1	−0.5	0.5	0.25	−0.25	0.25	0	1	−1	0.5	−0.5	0.25

m = mean; Aa = autosomal additive; Ad = autosomal dominance; AaAa = autosomal additive by additive epistasis; AaAd = autosomal additive by dominance epistasis; AdAd = autosomal dominance by dominance epistasis; Mea = additive maternal; Med = dominance maternal; Ca = additive cytotype; CaAa = additive cytotype by autosomal additive epistasis; and CaAd = additive cytotype by autosomal dominance epistasis.

crosses had been initiated, larvae from each cohort were picked over two subsequent days (one day for thermal stress traits and one day for body size traits) and placed into 18 replicate vials in total (nine per day) at a density of 50 larvae per vial. Once the larvae from each block were picked, six replicate vials per cohort, were placed at 18°C, 25°C and 29°C to develop, so that offspring could be measured for thermal stress resistance (Day 1, three replicate vials) and for all morphological traits (Day 2, three replicate vials) at each temperature.

CLIMATIC STRESS TRAITS

Heat resistance was scored as knock down time (Hoffmann et al. 2002), whereas cold resistance was scored as chill coma recovery time (David et al. 1998) and both were measured on females only. Females were separated from males under CO_2 anesthesia 48 h prior to stressing (thus females were assumed to have mated), and heat and cold resistance was measured on seven- and eight-day-old flies, respectively. For heat resistance, individual flies were placed in 10 mL dry vials and submerged in a water bath heated to 38.5°C and heat resistance was scored as the time taken (to the nearest second) for flies to be knocked down. For cold resistance, individual flies were placed in 10 mL dry vials and submerged in a water bath filled with 10% glycol solution and cooled to 0°C for 3 h. Chill coma recovery was assayed by scoring the time (to the nearest second) to recover (the ability to stand upright) at 25°C following a chill coma induced by a cold shock. Ten to 15 females per replicate vial were scored (total 30–45 females per cross). One-way analyses of variance (ANOVAs) showed no evidence for significant vial effects on cold or heat resistance at any temperature (data not shown).

MORPHOLOGICAL TRAITS

Wing centroid size and thorax length were measured on the same 10 males and 10 females from each of three replicate vials, which had developed at 18°C, 25°C, and 29°C. The right wing (or the left wing if the right was damaged) was removed from individual flies with fine forceps and mounted on glass slides with double-sided tape and protected with a cover slip. Wing images were captured with a Wild M3 dissector microscope attached to a digital camera and land-marked for the eight junctions of longitudinal veins with the wing margins or cross-veins (Liefting et al. 2009). Their x and y coordinates were recorded using the program TPSDIG version 1.31 written by F. J. Rohlf and wing size was calculated as centroid size, the square root of the sum of the squared inter-landmark distances (Hoffmann and Shirriffs 2002). Thorax length was measured as outlined in Hoffmann et al. (2007). To check for measurement error, we measured repeatability for both thorax and wing centroid size estimates. Repeat measures were found to be highly correlated for thorax ($r > 0.97$; $N = 50$) and wing centroid size ($r > 0.99$; $N = 100$). Furthermore, one-way ANOVAs found no evidence for significant vial effects on wing centroid or thorax size at any temperature (data not shown).

ANALYSIS

We used the following equations to explore the divergence in reaction norm average performance (elevation, eq. (1)), sensitivity

(slope, eq. (2)), and shape (curvature, eq. (3)) across the parental populations:

$$\text{Elevation} = \frac{\sum_1^n Z_i}{n}. \quad (1)$$

$$\text{Slope} = \frac{\sum_1^{n-1} S_i}{n-1};\ \ S_i = \frac{Z_{i+1} - Z_i}{T_{i+1} - T_i}. \quad (2)$$

$$\text{Curvature} = \frac{\sum_1^{n-2} C_i}{n-2};\ \ C_i = S_{i+1} - S_i. \quad (3)$$

For trait Z, n equals the number of test temperatures (3), and i represents the focal temperature (see Berger et al. 2014; Murren et al. 2014). Note that contrary to Murren et al. (2014), we have divided phenotypic differences by differences in temperatures (T) to compute reaction norm slopes and curvatures, because the widths of the temperature intervals were not identical (i.e., the difference between 29°C and 25°C was smaller than between 18°C and 25°C). These were calculated separately for each cohort, trait, sex, and species. As line-cross analyses examine the CGEs underlying population divergence in a phenotypic trait, we first examined divergence in reaction norm elevation, slope, and curvature between the parental populations using t tests. Plastic responses can include phenotypic changes with temperature, or underlying genetic/physiological changes that maintain the same trait value across temperature (canalization) (Debat and David 2001). As such, we also used a one-way ANOVA, with elevation, slope, and curvature as the dependent variable and cohort as a fixed effect, to look for any evidence of hybrid breakdown/heterosis of canalization across the cohorts. We then further examined the CGEs underlying divergence in the reaction norm parameters for traits where we observed either evidence for divergence across the parental lines, and/or across the cohorts (see next).

We used the software package SAGA (software for analysis of genetic architecture) in R (R Development Core Team 2014) to estimate the CGEs contributing to variation among cohorts for the reaction norm parameters for each trait in each sex of each species (Blackmon and Demuth 2016). SAGA uses a full information-theoretic approach that uses the finite sample size corrected version of the Akaike's information criterion (AICc) to explore all possible models and make unbiased and, when appropriate, model-averaged estimates of the contribution of CGEs to cohort means. This approach has the advantage of assessing the potential model space, quantifying model selection uncertainty, and using model weighted averaging to accurately estimate CGEs. Traditional line-cross analysis (e.g., joint scaling tests, Lynch and Walsh 1998) depend on identifying the best model and interpreting the CGEs that are included in that model. SAGA is able to obtain accurate estimates of the CGEs that are not dependent on the ability to specify one overall model as best. If the Akaike weight (wi) of the best model is 0.95 or greater, SAGA will perform parameter estimation under a single model. If no model reaches this threshold, then SAGA constructs a 95% confidence set of models that contain the minimum number of models whose wi sum to 0.95. SAGA then computes model-averaged results for the 95% confidence set and provides estimates of variable importance (vi) calculated by summing wi of all models in which a CGE occurs. The vi score provides evidence that a CGE is important even if its contribution is small or poorly defined. Although one of the strengths of SAGA is that it alleviates the use of strict arbitrarily defined P values, Blackmon and Demuth (2016) generally found CGEs with a vi score greater than 0.5 were most likely to be included in the model containing the 95% confidence model set.

To directly compare models with different CGEs for elevation, slope, and curvature, we calculated AICc using the following equation (Burnham and Anderson 2002), where n is the number of cohorts and K is the number of parameters being estimated:

$$\text{AICc} = \text{AIC} + \frac{2K(K+1)}{n-K-1}. \quad (4)$$

We first calculated AICc for the model that contained all the CGEs with a $vi > 0.5$ (or 0.3 if there was high model uncertainty and no CGEs had a $vi > 0.5$) for each reaction norm parameter (e.g., elevation). We then compared that AICc score to the model containing the CGEs with a $vi > 0.5$ for the other reaction norm parameters for that trait. For example, for elevation in heat knock down in *D. melanogaster*, we compared the AICc for the model containing the CGEs with a $vi > 0.5$ for elevation to the AICc scores for the models containing the CGEs with a $vi > 0.5$ for slope and curvature, to ask whether the CGEs identified as important for elevation are a better fit than those identified as important for the other reaction norm parameters.

Results

The mean elevation, slope, and curvature of all 14 cohorts for each interpopulation cross for each trait in each species are shown in Figures S1–S3. We found significant divergence between the parental populations and/or across the cohorts for trait mean—elevation and plasticity—slope, and curvature for all traits (Table S1; Figs. 1 and 2). Thus, we further examined the CGEs underlying clinal divergence in elevation, slope, and curvature for all traits in both species. Similar to Blackmon and Demuth (2016), no single model had a wi sufficient to ignore model selection uncertainty (Akaike weight [wi] of the best model < 0.95) for elevation, slope, or curvature for any trait (Table S2), so for all traits, we used the model-averaged results for the 95% confidence set to get estimates of the CGEs underlying clinal divergence. In general, the lower the wi of the best model and or the more models contained in the 95% confidence set, the greater the degree of model selection uncertainty (Blackmon and Demuth 2016). For a

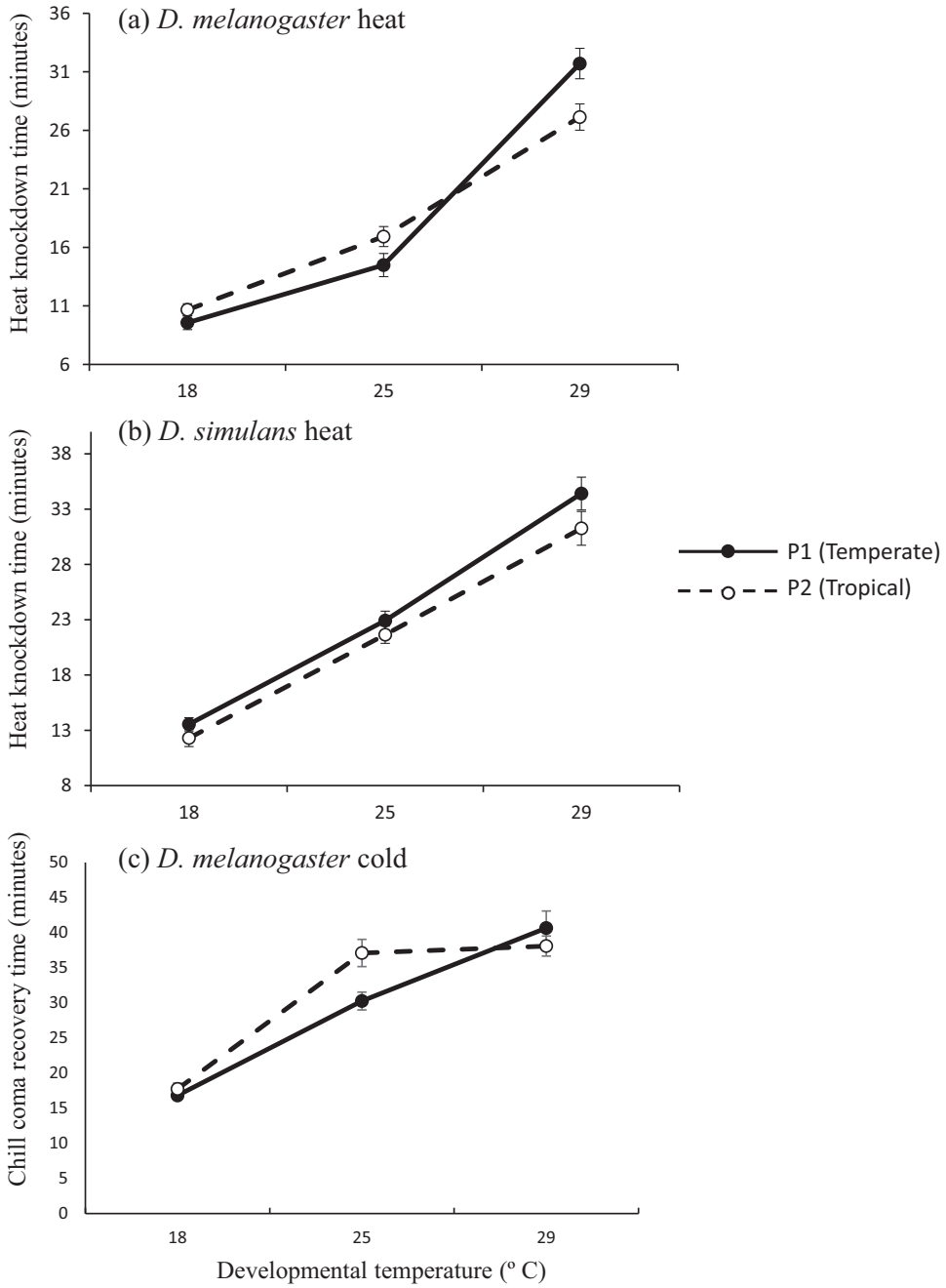

Figure 1. **Thermal reaction norm of mean heat knockdown time at 38.5°C in (A) *Drosophila melanogaster* females and (B) *Drosophila simulans* females, and mean chill coma recovery time at 25°C after 3 h at 0°C in (C) *D. melanogaster* females under different developmental/adult acclimation temperatures. Error bars are 1 SE.**

small number of traits, there was high model selection uncertainty for some of the reaction norm parameters, evidenced by low *wi* and the high number of models required (Table S2). However, in these situations it is still possible to infer which CGEs are likely to be important, albeit with less confidence in estimating their true magnitude or sign because they depend on the other components in the model and SEs will therefore overlap zero (Blackmon and Demuth 2016).

PATTERNS IN THE GENETIC ARCHITECTURE UNDERLYING CLINAL DIVERGENCE IN ELEVATION, SLOPE, AND CURVATURE OF STRESS AND MORPHOLOGICAL TRAITS

Across all traits, autosomal additive effects contributed to clinal divergence in less than a third of crosses, and were detected more frequently for divergence in mean performance—elevation (observed in 55% of crosses), than for plasticity—slope (observed in

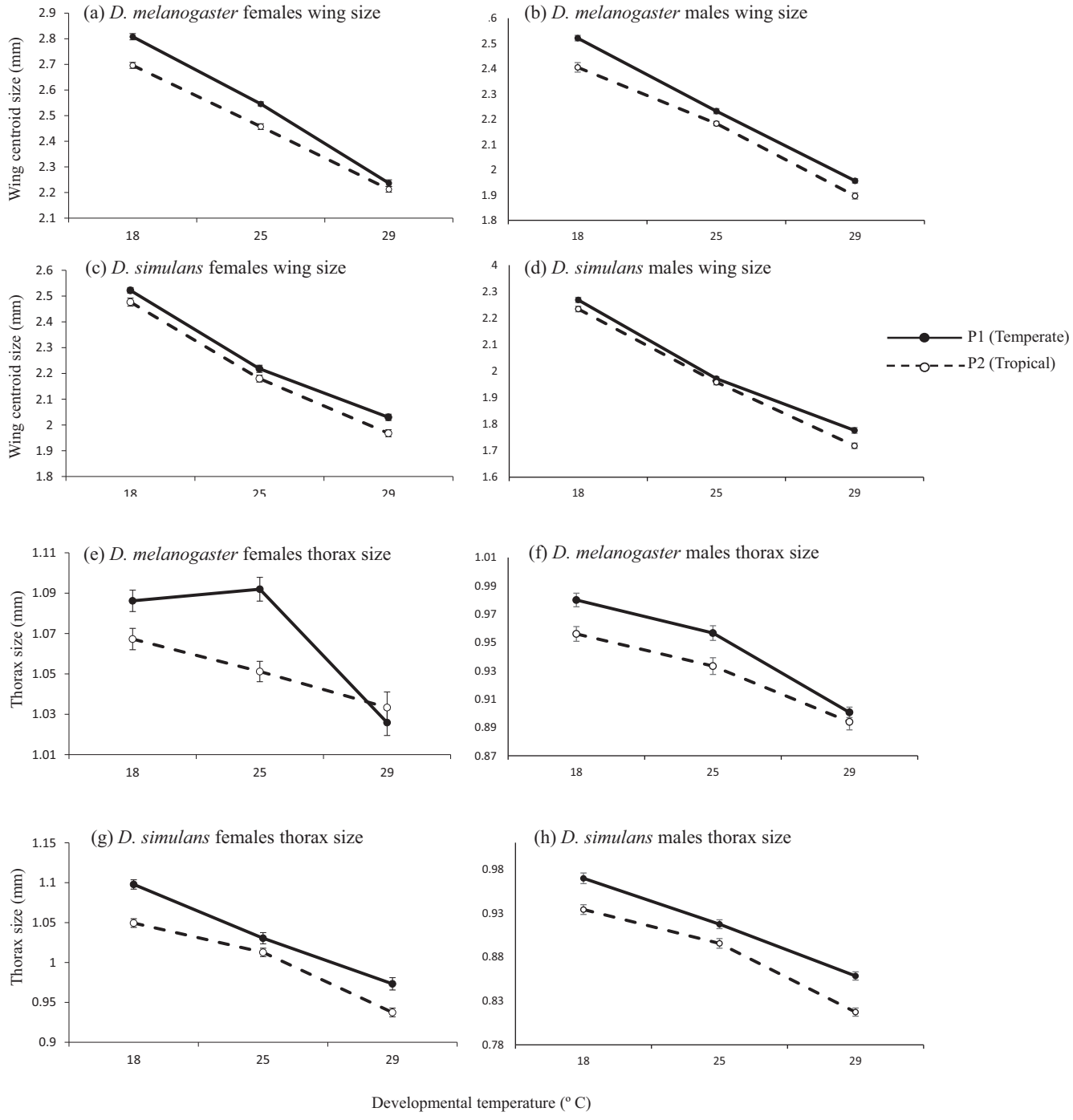

Figure 2. Thermal reaction norm for mean wing centroid size in *Drosophila melanogaster* (A) females and (B) males, mean wing centroid size in *Drosophila simulans* (C) females and (D) males; mean thorax size in *D. melanogaster* (E) females and (F) males; and mean thorax size in *D. simulans* (G) females and (H) males. Error bars are 1 SE.

9% of crosses), or curvature (observed in 18%) (Table 2). Autosomal additive effects were important for divergence in elevation for heat knockdown time in *D. simulans*, wing size in *D. melanogaster* females and males, wing size in *D. simulans* females, thorax size in *D. melanogaster* males, and thorax size in *D. simulans* females (Figs. 3, 4, and 5). For slope, autosomal additive effects were only important for divergence in thorax size in *D. melanogaster* females, and for curvature, autosomal additive effects contributed to divergence in cold in *D. melanogaster* females, and wing size in female *D. melanogaster* (Figs. 3, 4, and 5).

Autosomal dominance effects were less common than autosomal additive effects, contributing to clinal divergence in

Table 2. Summary of the CGEs underlying divergence in elevation, slope, and curvature (averaged across all traits/sex/ species) in the stress and morphological traits (averaged across sex/ species) and overall.

	N	Additive		Dominance		Autosomal epistasis		Cytotype by autosomal epistasis		Epistasis overall		Cytotype		Maternal	
Elevation	11	6	(55%)	3	(27%)	5	(45%)	2	(18%)	5	(45%)	2	(18%)	5	(45%)
Slope	11	1	(9%)	1	(9%)	5	(45%)	7	(64%)	10	(91%)	4	(36%)	8	(73%)
Curvature	11	2	(18%)	3	(27%)	9	(82%)	5	(45%)	10	(91%)	2	(18%)	5	(45%)
Stress	9	2	(22%)	3	(33%)	7	(78%)	7	(78%)	8	(89%)	2	(22%)	6	(67%)
Morph	24	7	(29%)	4	(17%)	12	(50%)	7	(29%)	17	(71%)	6	(25%)	12	(50%)
Overall	33	9	(27%)	7	(21%)	19	(58%)	14	(42%)	25	(76%)	8	(24%)	18	(55%)

The discrete values are the number of crosses where these effects were found to be important (*vi* > 0.5, or *vi* > 0.3, where there was high model selection uncertainty) and the percentage is the fraction of all crosses in a particular category where these effects were observed.

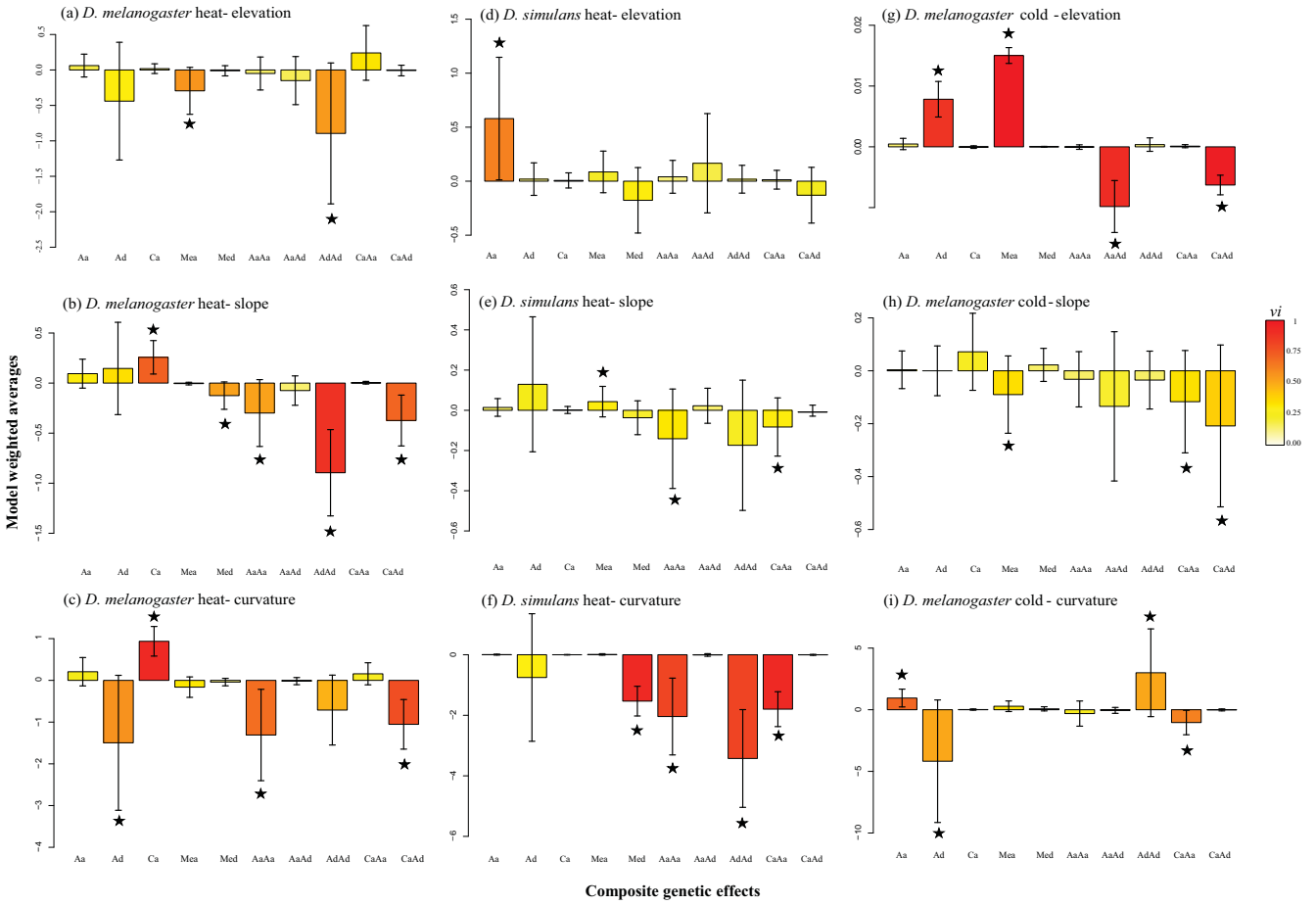

Figure 3. Model-weighted average values for the different CGEs contributing to the divergence between tropical and temperate populations in reaction norm elevation, slope, and curvature for heat knock down time in *Drosophila melanogaster* (A, B, C) and *Drosophila simulans* (E, F, G), and chill coma recovery time in *D. melanogaster* (H, I, J). Bars are colored based on *vi* (variable importance) score, which provides evidence that a CGE is important even if its contribution is small or poorly defined. Error bars indicate the unconditional SEs and stars indicate which CGEs are compared across the different reaction norm parameters. The direction of each CGE indicates whether the average effect across all loci is positive or negative in relation to the reference population, P1. m = mean; Aa = autosomal additive; Ad = autosomal dominance; AaAa = autosomal additive by additive epistasis; AaAd = autosomal additive by dominance epistasis; AdAd = autosomal dominance by dominance epistasis; Mea = additive maternal; Med = dominance maternal; Ca = additive cytotype; CaAa = additive cytotype by autosomal additive epistasis; and CaAd = additive cytotype by autosomal dominance epistasis.

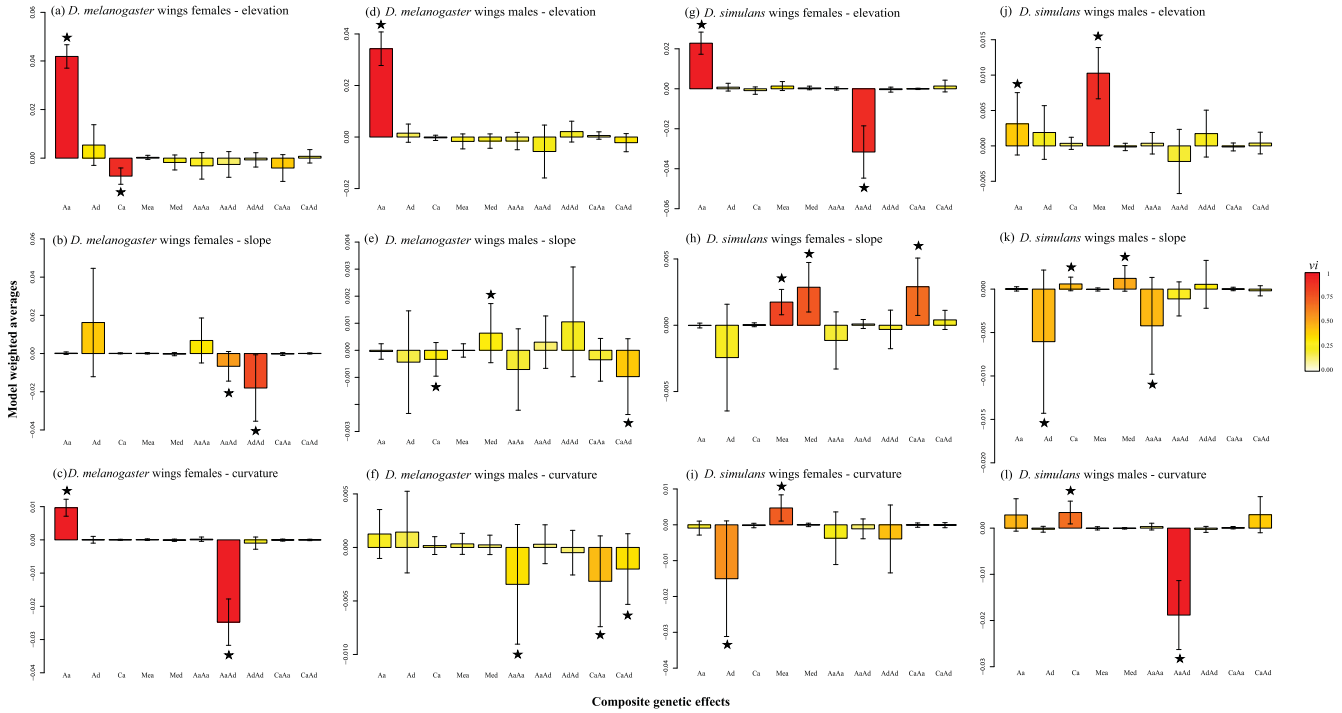

Figure 4. Model-weighted average values for the different CGEs contributing to the divergence between tropical and temperate populations in reaction norm elevation, slope, and curvature for wing centroid size in *D. melanogaster* females (A, B, C) and males (D, E, F), and *D. simulans* females (G, H) and males (I, J, K). Bars are colored based on *vi* (variable importance) score, which provides evidence that a CGE is important even if its contribution is small or poorly defined. Error bars indicate the unconditional SEs and stars indicate which CGEs are compared across the different reaction norm parameters/traits/sex. The direction of each CGE indicates whether the average effect across all loci is positive or negative in relation to the reference population, P1. m = mean; Aa = autosomal additive; Ad = autosomal dominance; AaAa = autosomal additive by additive epistasis; AaAd = autosomal additive by dominance epistasis; AdAd = autosomal dominance by dominance epistasis; Mea = additive maternal; Med = dominance maternal; Ca = additive cytotype; CaAa = additive cytotype by autosomal additive epistasis; and CaAd = additive cytotype by autosomal dominance epistasis.

elevation, slope, and/or curvature in only 21% of crosses overall (Table 2). Similar to autosomal additive effects, autosomal dominance effects were observed more often for divergence in elevation (27% of crosses), than in slope (9% of crosses), but were observed in a similar frequency for curvature (27%) (Table 2). Autosomal dominance effects were important for divergence in curvature for heat knock down time in *D. melanogaster*, divergence in elevation and curvature for cold in *D. melanogaster*, divergence in curvature and elevation for wing size in *D. simulans* females and males, respectively, and divergence in thorax size in *D. melanogaster* and *D. simulans* males (Figs. 3, 4, and 5).

Autosomal epistatic and cytotype by autosomal epistatic effects were found more frequently than autosomal additive or dominance effects, detected in 58 and 42% of crosses, respectively (Table 2). Autosomal epistatic effects were more common for curvature (found in 82% of crosses) than for elevation (45%) or slope (45%), whereas cytotype epistatic effects were more common for both measures of plasticity (slope: 64%, curvature: 45%), than for elevation (18%) (Table 2). Autosomal additive by additive epistatic effects were only important for divergence in plasticity, detected for slope and curvature in female heat knockdown time in *D. melanogaster* and *D. simulans* females, slope for wing size in *D. simulans* males, and curvature in wing size and thorax size in *D. melanogaster* males (Figs. 3, 4, and 5). Autosomal additive by dominance and/or autosomal dominance by dominance epistasis were important for elevation for heat and cold resistance in *D. melanogaster*, wing size in *D. simulans* females, and thorax size in *D. simulans* females and males; slope for heat resistance and wing and thorax size in *D. melanogaster* females; and curvature in heat resistance and thorax size in *D. simulans* females, wing size in *D. simulans* males and cold resistance, wing and thorax size in *D. melanogaster* females (Figs. 3, 4, and 5).

Maternal effects were also common for divergence in elevation, slope, and curvature and were observed in just over half of all crosses, whereas cytotype effects were only observed in 24% of crosses overall. Maternal and cytotype effects contributed more frequently to divergence in slope (73 and 36%, respectively) than for elevation (45 and 18%, respectively) or curvature (45 and 18%, respectively) (Table 2).

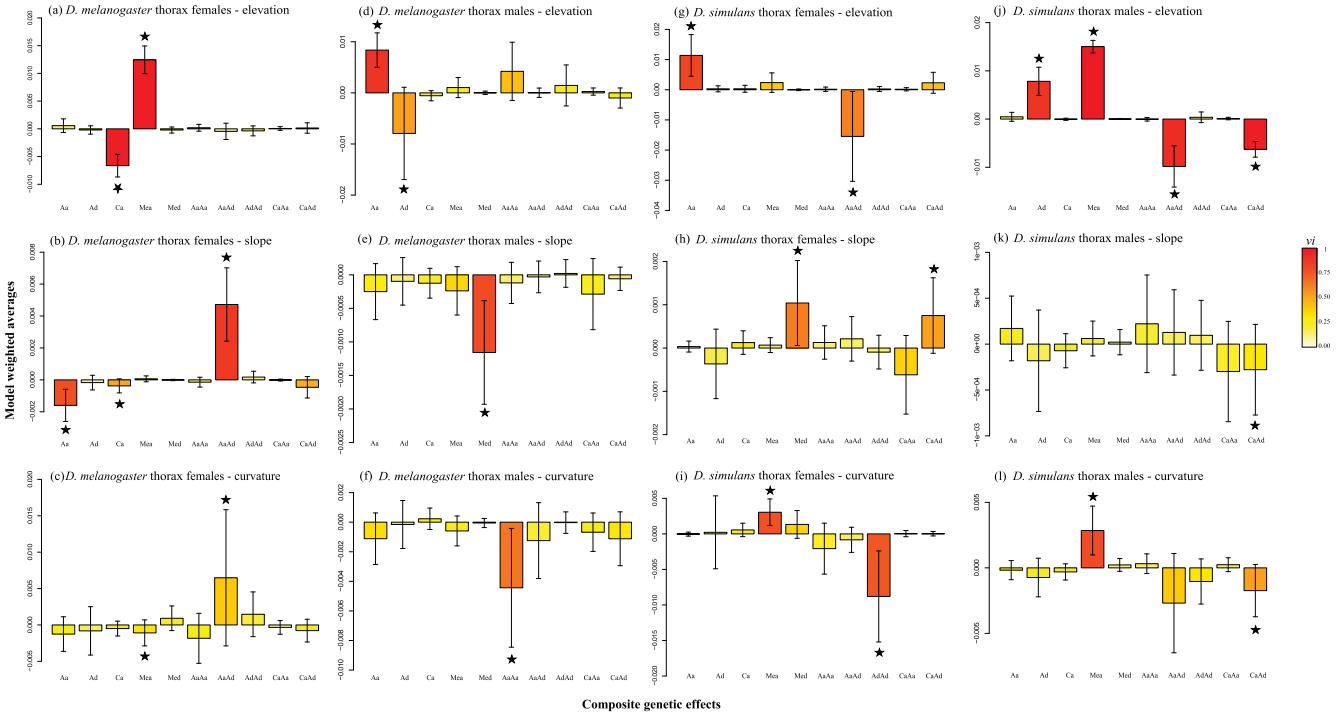

Figure 5. **Model-weighted average values for the different CGEs contributing to the divergence between tropical and temperate populations in reaction norm elevation, slope, and curvature for thorax size of *Drosophila melanogaster* female (a, b, c) and males (d, e), and *Drosophila simulans* females (f, g, h) and males (i, j). Bars are colored based on v*i* (variable importance) score, which provides evidence that a CGE is important even if its contribution is small or poorly defined. Error bars indicate the unconditional SEs and stars indicate which CGEs are compared across the different reaction norm parameters. The direction of each CGE indicates whether the average effect across all loci is positive or negative in relation to the reference population, P1. m, mean; Aa, autosomal additive; Ad, autosomal dominance; AaAa, autosomal additive by additive epistasis; AaAd, autosomal additive by dominance epistasis; AdAd, autosomal dominance by dominance epistasis; Mea, additive maternal; Med, dominance maternal; Ca, additive cytotype; CaAa, additive cytotype by autosomal additive epistasis; CaAd, additive cytotype by autosomal dominance epistasis.**

Across all reaction norm parameters, the frequency of additive effects was similar for stress and morphological traits (Table 2). Dominance, autosomal, and cytotype epistasis, and maternal effects were more common for divergence in the stress traits, while cytotype effects were slightly more common for divergence in the morphological traits (Table 2).

DIFFERENCES IN GENETIC ARCHITECTURE BETWEEN ELEVATION, SLOPE, AND CURVATURE

Generally, different CGEs were identified as important for explaining clinal divergence in the different reaction norm parameters (i.e., elevation, slope, and curvature) for each trait (Figs. 3, 4, and 5). Indeed, when we compared the AICc scores for the models including CGEs with a variable importance (*vi*) of more than 0.5 (or $vi > 0.3$ if there was high model selection uncertainty) for each reaction norm parameter (e.g., elevation) against AICs scores that included the same CGEs identified for the other reaction norm parameters (e.g. slope and curvature) for each trait, we generally found strong support that the models that best described divergence in elevation, slope, and curvature were different (AICc best model $>$ 2 AICc than the alternate model) (Tables S3–S5), suggesting that the genetic basis underlying divergence in these reaction norm parameter values differ. The only traits where this was not the case were slope (vs. elevation and curvature) and curvature (vs. elevation and slope) for wing size in *D. melanogaster* males, slope (vs. curvature) for wing size in *D. simulans* males (Table S4), and elevation (vs. curvature) for thorax size in *D. simulans* females, and slope (vs. curvature) and curvature and elevation for thorax size in *D. simulans* males (Table S5).

We also found little evidence that the genetic basis underlying divergence in elevation, slope, and curvature is similar across the sexes. With the exception of curvature for thorax size in *D. simulans* and *D. melanogaster*, and elevation and slope for wing size in *D. melanogaster*, the models that best described divergence in elevation, slope, and curvature differed across sexes in both species (AICc best model for females $>$ 2 AICc than the alternate model for males and vice versa) (Tables S6 and S7). Additionally, we found little evidence that the genetic basis underlying divergence in elevation, slope, and curvature is similar across species. With the exception of slope for wing and thorax size in males,

and curvature in thorax size in females, all of the models that best described divergence in elevation, slope, and curvature differed across *D. melanogaster* and *D. simulans* (AICc best model for *D. melanogaster* > 2 AICc than the alternate model for *D. simulans* and vice versa) (Tables S8–S10).

Discussion

The genetic basis of phenotypic plasticity is not well understood, despite being important for modeling and understanding the evolution of plastic responses in nature. The extent to which additive or nonadditive effects contribute to the evolution of plastic responses is still not clear even though physiological epistasis (i.e., epistasis between regulatory and structural loci) is central to evolutionary models of plasticity (including the gene regulation model) and canalization (Scheiner 1993; Wagner et al. 1997; Rice 1998; Flatt 2005). We provide the first detailed examination of the quantitative genetic basis of divergence in mean thermal response and thermal plasticity in locally adapted outbred populations of *D. melanogaster* and *D. simulans*. Our analysis suggests that the genetic basis of population divergence in reaction norm elevation, slope, and curvature differs, and that epistasis frequently contributes to divergence in plasticity. We discuss the implication of these findings below.

There has been considerable interest in the extent to which epistasis may contribute to the adaptive divergence of key traits (Wright 1931; Whitlock et al. 1995; Fenster et al. 1997; Wade and Goodnight 1998), as well as its role in the maintenance of genetic variance and the rate of evolution (Hansen 2013). Using crosses between diverged populations or lines/stocks, several studies have shown that epistasis is indeed important for adaptive divergence in morphological, stress, and fitness traits (Hard et al. 1992; Armbruster et al. 1997; Gilchrist and Partridge 1999; Schiffer et al. 2006; van Heerwaarden and Sgro 2011). However, only a small number of studies have used this approach to examine the quantitative genetic basis of plasticity (e.g., Perkins and Jinks 1973; Connolly and Jinks 1975; Pooni et al. 1987). These studies found evidence that dominance and epistasis contribute to both mean performance and environmental sensitivity (plasticity), suggesting that nonadditive effects may commonly underlie the evolution of plasticity. Nonetheless, these studies neither considered maternal effects, which may contribute significantly to F_2 breakdown, nor did they use a comprehensive crossing design that allowed the partitioning of different epistatic effects. Furthermore, no such studies have been performed to understand the genetic basis of divergence in plasticity in outbred populations originating from, and adapted to, different biogeographical habitats. Although we used flies that had been in the laboratory for six to nine generations before starting the experiment, and may thus have undergone some level of laboratory adaptation (Santos et al. 2012), this is the first study to our knowledge that has used a comprehensive crossing design to examine the genetic architecture underlying plasticity using locally adapted outbred populations recently collected from different biogeographical locations. Consistent with these earlier studies (e.g. Perkins and Jinks 1973; Connolly and Jinks 1975; Pooni et al. 1987), overall we observed that additive, dominance, epistatic, and maternal effects all contributed to divergence in the mean (elevation) and plasticity (slope and curvature) of body size and/or thermal tolerance in *D. melanogaster* and *D. simulans*. Thus, both additive and nonadditive genetic effects underlie the evolution of the mean of these traits (van Heerwaarden and Sgrò 2011) and their plastic responses to temperature (this study).

In both species, we also observed that the relative contribution of nonadditive genetic effects underlying the divergence in trait mean (elevation) and thermal plasticity (slope and curvature) differed. Autosomal additive effects were more common for divergence in mean values (elevation), than for plasticity (slope or curvature), whereas autosomal dominance effects contributed more frequently to divergence in elevation and curvature than for slope. Cytotype and maternal effects were detected more frequently to divergence in slope than elevation or curvature. Importantly, epistatic effects were more common for plasticity than trait mean (elevation), consistent with Connolly and Jinks (1975). Across all traits, autosomal and/or cytotype epistasis was observed in 91% of crosses for both slope and curvature (cf. to 45% of crosses for elevation) (Table 2), suggesting that epistasis may be crucial to the evolution of plastic responses. This finding has important implications for genetic models of plasticity, which generally ignore nonadditive genetic variation (Via et al. 1995; Berrigan and Scheiner 2004). It is also important for understanding and predicting how traits and their plasticity evolve (Carter et al. 2005; Hallander and Waldmann 2007). This is because studies have shown that epistasis can alter additive genetic variance of traits under selection, and some models suggest that epistasis may accelerate or constrain evolutionary responses to selection (Carter et al. 2005; Carlborg et al. 2006; Hallander and Waldmann 2007). Specifically, positive epistasis, where genes tend to reinforce each other's effects in the direction of selection, will increase additive genetic variance and accelerate the response to selection, while negative epistasis, where genes tend to diminish each other's effects in the direction of selection and reduce additive genetic variance, will reduce the response (Hansen and Wagner 2001). Thus, if epistasis is pervasive, predicted evolutionary responses in traits or their plasticity based on estimates of additive genetic variance alone may be inaccurate.

In addition to differences between the genetic basis of mean (elevation) and plasticity (slope), we also found that the quantitative genetic basis of the two reaction norm parameters that describe plasticity (slope and curvature) differs. Although quantitative genetic models have not explored the effect of additive

and nonadditive genetic effects on the evolution of plasticity separately, de Jong and Gavrilets (2000) predicted that the additive genetic variance for the elevation and slope of a linear reaction norm, as well as their covariance, should depend on the level of variation in their developing environment, decreasing with increasing variation in the environment of development. Furthermore, Lande (2009) found that after large and sudden changes in environment exceeding typical background environmental fluctuations, the proportion of additive genetic variance of a trait in the new environment may increase due to additive genetic variance in plasticity. Given that plasticity is predicted to evolve when populations experience spatial and/or temporal environmental heterogeneity (Via and Lande 1985; Gabriel and Lynch 1992; Gabriel et al. 2005), environmental variability may directly influence the evolution of plasticity. Consistent with this prediction, de Jong and Gavrilets (2000) showed that the additive genetic variance for elevation and slope (and their genetic covariance) of morphological traits in *Drosophila* both decrease with increasing variation in the environment of development. Furthermore, as long as some loci that influence only the slope are present (i.e., pleiotropy is not complete, as predicted under the gene regulation hypothesis), the genetic variance in slope (plasticity) is predicted to decrease faster with increased environmental variance than the genetic variance in elevation (mean) (de Jong and Gavrilets 2000). We have only measured the relative contributions of different genetic effects to population divergence, rather than assessing standing additive and nonadditive genetic variance, and are thus unable to directly compare our results to those of de Jong and Gavrilets (2000). However, if the additive genetic variance in elevation and slope (and their genetic covariance) does indeed change across environment as predicted (de Jong and Gavrilets 2000), then the differences in the overall quantitative genetic architecture (additive, dominance, epistasis, maternal, etc.) for mean trait values (elevation) and plasticity (slope and curvature) between populations originating from, and locally adapted to, different environments as observed in our study may be expected.

When exploring the level of evolutionary divergence in different reaction norm parameters across diverged populations and species, Murren et al. (2014) found that divergence in plasticity (slope and curvature) among closely related species were greater than divergence in trait means (elevation), indicating that microevolutionary changes in plasticity may be more common than evolved shifts in trait means. They also found that differences in curvature among closely related species were greater than differences in slope. Higher levels of divergence in plasticity (slope and curvature) than elevation contrast with quantitative genetic experiments that have shown that heritability for plasticity is lower than for elevation (Scheider 1993), but are consistent with theoretical models, which suggest that plasticity may increase with large and sudden changes in environment, with relatively little change to the additive genetic variance of the trait in the original environment (Lande 2009). These results suggest that there is significant potential for the additive genetic variance and selection responses to differ for trait mean (elevation) and plasticity (slope and curvature). Our study, which revealed differences in genetic architecture for divergence in trait mean and plasticity, are consistent with such predictions. In addition, our results are consistent with models that suggest that epistasis may play an important role in evolutionary shifts in traits and plasticity (Carter et al. 2005).

We also found that the genetic effects underlying divergence in plasticity for heat/cold resistance and body size generally differed, suggesting that the genetic basis of plasticity may be trait-specific. Given the close association between body and environmental temperature in ectotherms (Cossins and Bowler 1987), resistance to temperature extremes is likely to be important for surviving and adapting to different thermal environments, as well as dictating species distributions (Sunday et al. 2011; Overgaard et al. 2014). Overall, we found that dominance, autosomal, and cytotype epistasis and maternal effects were slightly more common for divergence in the plasticity of heat/cold resistance compared to size. Roff and Emerson (2006) hypothesized that fitness traits would show higher levels of epistasis and dominance because these traits should be subjected to intense natural selection, which is predicted to deplete additive genetic variation, and leave segregating loci with primarily dominance and epistatic effects. Indeed, in a meta-analysis of line-cross studies, they observed more epistasis for fitness traits than morphological traits (Roff and Emerson 2006). However, in comparisons between tropical and temperate populations of *Drosophila*, the frequency of nonadditive genetic effects was higher for morphological traits than stress traits (Gilchrist and Partridge 1999; Kennington et al. 2001; van Heerwaarden and Sgro 2011). Our data suggest that different genetic architectures underpin the mean and plasticity of morphological and climatic stress resistance traits.

Finally, the genetic effects underlying population divergence in the different reaction norm parameters differed between *D. melanogaster* and *D. simulans*, and across the sexes. Despite sharing similar distributions and thus similar environmental conditions, these results suggest that similar selective forces can cause divergence in traits and their plasticity via quite different types of gene action and interaction. These results are also consistent with evolutionary patterns observed in a recent meta-analysis (Murren et al. 2014), which showed that divergence in reaction norm shape varies between organisms, traits, and environments. Taken together, these findings suggest that the evolution of plasticity is complex, and that this complexity must be considered in future studies of plasticity. Although we only examined one population from each latitude for each species, which may limit our ability to make general statements about which particular genetic effects underlie clinal divergence for each trait specifically, our

replication at the species level supports our general findings that different genetic effects underlie mean performance and plasticity, and that epistasis contributes frequently to evolutionary divergence in plasticity more broadly. In addition, the fact that independent studies on *D. melanogaster* have repeatedly shown clinal patterns in the traits we examined (e.g., heat: Hoffmann et al. 2002; Sgro et al. 2010; Cockerell et al. 2014; size: James et al. 1995; 1997; van Heerwaarden and Sgro 2011; cold: Hoffmann et al. 2002, 2005), that also reflect the level of population divergence we see in this study, suggests that the divergence in these traits and their plasticity are likely due to local adaptation.

In conclusion, the genetic basis of geographic divergence in mean performance and both measures of plasticity differs in both *D. melanogaster* and *D. simulans*. In addition, the prevalence of nonadditive genetic effects for plasticity, as well as trait means (van Heerwaarden and Sgro 2011) suggests that nonadditive genetic effects (particularly epistasis) are important for the evolution of both trait means and trait plasticity. Current evolutionary models of adaptive responses to environmental change largely assume that additive genetic effects will be important for the evolution of both trait mean and plasticity (Fisher 1930; de Jong 1995; Coyne et al. 1997, 2000; Hill et al. 2008). The extent to which predictions of evolutionary shifts in plasticity and trait means will differ if nonadditive effects are explicitly considered remains to be assessed.

AUTHOR CONTRIBUTIONS

C.M.S. and B.V.H. designed the experiments. B.V.H. performed the experiments and analyzed the data. Both authors contributed to the final manuscript.

ACKNOWLEDGMENTS

We thank the Australian Research Council, the Commonwealth Environment Research Facility, the Science and Industry Endowment Fund, and Monash University for financial support. We would also like to thank V. Kellermann, F. Cockerell, R. Lee, J. Jory, and W. Yee for technical support.

DATA ARCHIVING

The doi for our data ihttps://doi.org/10.5061/dryad.43j31.

LITERATURE CITED

Armbruster, P., W. E. Bradshaw, and C. M. Holzapfel. 1997. Evolution of the genetic architecture underlying fitness in the pitcher-plant mosquito, *Wyeomyia smithii*. Evolution 51:451–458.

Atkinson, D. 1994. Temperature and organism size—a biological law for ectotherms. Pp. 1–58 *in* M. Begon and A. H. Fitter, eds. Advances in ecological research, Vol. 25. Academic Press, London, UK.

Azevedo, R. B. R., A. C. James, J. McCabe, and L. Partridge. 1998. Latitudinal variation of wing: thorax size ratio and wing-aspect ratio in *Drosophila melanogaster*. Evolution 52:1353–1362.

Barnes, P. T., B. Holland, and V. Courreges. 1989. Genotype-by-environment and epistatic interactions in *Drosophila melanogaster*—the effects of GPDH allozymes, genetic background and rearing temperature on larval developmental time and viability. Genetics 122:859–868.

Barton, N. H. 1995. A general-model for the evolution of recombination. Genet. Res. 65:123–144.

Berger, D., R. J. Walters, and W. U. Blanckenhorn. 2014. Experimental evolution for generalists and specialists reveals multivariate genetic constraints on thermal reaction norms. J. Evol. Biol. 27:1975–1989.

Berrigan, D., and S. M. Scheiner. 2004. Modeling the evolution of phenotypic plasticity. Pp. 82–97 *in* T. J. Dewitt, and S. M. Scheiner, eds. Phenotypic plasticity: functional and conceptual approaches. Oxford Univ. Press, New York.

Blackmon, H. and J. P. Demuth. 2016. An information-theoretic approach to estimating the composite genetic effects contributing to variation among generation means: moving beyond the joint-scaling test for line cross analysis. Evolution 70:420–432.

Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. Adv. Genet. 13:115–155.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York.

Carlborg, O., L. Jacobsson, P. Ahgren, P. Siegel, and L. Andersson. 2006. Epistasis and the release of genetic variation during long-term selection. Nat. Genet. 38:418–420.

Carson, H. L. and A. R. Templeton. 1984. Genetic revolutions in relation to speciation phenomena—the founding of new populations. Annu. Rev. Ecol. Syst. 15:97–131.

Carter, A. J. R., J. Hermisson, and T. F. Hansen. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theoret. Pop. Biol. 68:179–196.

Chevin, L. M., R. Gallet, R. Gomulkiewicz, R. D. Holt, and S. Fellous. 2013. Phenotypic plasticity in evolutionary rescue experiments. Phil. Trans. R. Soc. B. 36820120089.

Chown, S. L., and K. J. Gaston. 2010. Body size variation in insects: a macroecological perspective. Biol. Rev. 85:139–169.

Cockerell, F. E., C. M. Sgro, and S. W. McKechnie. 2014. Latitudinal dines in heat tolerance, protein synthesis rate and transcript level of a candidate gene in *Drosophila melanogaster*. J. Insect Physiol. 60:136–144.

Connolly, V., and J. L. Jinks. 1975. Genetic architecture of general and specific environmental sensitivity. Heredity 35:249–259.

Cossins, A. R., and K. Bowler. 1987. Temperature biology of animals. Chapman & Hall, New York.

Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: a critique of Sewall Wright's shifting balance theory of evolution. Evolution 51:643–671.

———. 2000. Is Wright's shifting balance process important in evolution? Evolution 54:306–317.

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. Therm. Biol. 22:441–451.

David, R. J., P. Gibert, E. Pla, G. Petavy, D. Karan, and B. Moreteau. 1998. Cold stress tolerance in Drosophila: analysis of chill coma recovery in *D. melanogaster*. J. Therm. Biol. 23:291–299.

Debat, V., and P. David. 2001. Mapping phenotypes: canalization, plasticity and developmental instability. Trends Ecol. Evol. 16:555–561.

Des Marais, D. L., K. M. Hernandez, and T. E. Juenger. 2013. Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. Annu. Rev. Ecol. Evol. Syst. 44:5–29.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

Falconer, D. S. 1952. The problem of environment and selection. Am. Nat. 86:293–298.

Fallis, L. C., J. J. Fanara, and T. J. Morgan. 2014. Developmental thermal plasticity among *Drosophila melanogaster* populations. J. Evol. Biol. 27:557–564.

Fenster, C. B., L. F. Galloway, and L. Chao. 1997. Epistasis and its consequences for the evolution of natural populations. Trends Ecol. Evol. 12:282–286.

Fethi, B., C. Hanbary, and E. G. Mohamed. 2011. Genetic adaptability of inheritance of resistance to biotic and abiotic stress level on crop: role of epistasis. Afr. J. Biotechnol. 10:19913–19917.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, U.K.

Flatt, T. 2005. The evolutionary genetics of canalization. Q. Rev. Biol. 80:287–316.

Fox, C. W., M. E. Czesak, and W. G. Wallin. 2004. Complex genetic architecture of population differences in adult lifespan of a beetle: nonadditive inheritance, gender differences, body size and a large maternal effect. J. Evol. Biol. 17:1007–1017.

Gabriel W., and M. Lynch. 1992. The selective advantage of reaction norms for environmental tolerance. J. Evol. Biol. 5:41–59.

Gabriel W., B. Luttbeg, A. Sih, and R. Tollrian. 2005. Environmental tolerance, heterogeneity, and the evolution of reversible plastic responses. Am. Nat. 166:339–353.

Ghalambor, C. K., R. B. Huey, and P. R. Martin. 2004. Why mountain passes are higher in the tropics - Revisited. Integr. Comp. Biol. 44: 558–558.

Gilchrist, A. S. and L. Partridge. 1999. A comparison of the genetic basis of wing size divergence in three parallel body size clines of *Drosophila melanogaster*. Genetics 153:1775–1787.

Gimelfarb, A. 1989. Genotypic variation for a quantitative character maintained under stabilizing selection without mutations—epistasis. Genetics 123:217–227.

Hallander, J., and P. Waldmann. 2007. The effect of non-additive genetic interactions on selection in multi-locus genetic models. Heredity 98:349–359.

Hansen, T. F. 2013. Why epistasis is important for selection and adaptation. Evolution 67:3501–3511.

Hansen, T. F., and G. P. Wagner. 2001. Modeling genetic architecture: a multilinear theory of gene interaction. Theor. Pop. Biol. 59:61–86.

Hard, J. J., W. E. Bradshaw, and C. M. Holzapfel. 1992. Epistasis and the genetic-divergence of photoperiodism between populations of the pitcher-plant mosquito, *Wyeomyia smithii*. Genetics 131:389–396.

Hermisson, J., T. F. Hansen, and G. P. Wagner. 2003. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. Am. Nat. 161:708–734.

Hill, W. G., M. E. Goddard, and P. M. Visscher. 2008. Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet. 4:e1000008.

Hoffmann, A. A., and J. Shirriffs. 2002. Geographic variation for wing shape in *Drosophila serrata*. Evolution 56:1068–1073.

Hoffmann, A. A., A. Anderson, and R. Hallas. 2002. Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. Ecol. Lett. 5:614–618.

Hoffmann, A. A., J. Shirriffs, and M. Scott. 2005. Relative importance of plastic vs. genetic factors in adaptive differentiation: geographical variation for stress resistance in *Drosophila melanogaster* from eastern Australia. Funct. Ecol. 19:222–227.

Hoffmann, A. A., E. Ratna, C. M. Sgro, M. Barton, M. Blacket, R. Hallas, S. De Garis, and A. R. Weeks. 2007. Antagonistic selection between adult thorax and wing size in field released *Drosophila melanogaster* independent of thermal conditions. J. Evol. Biol. 20:2219–2227.

Huey, R. B., and D. Berrigan. 2001. Temperature, demography, and ectotherm fitness. Am. Nat. 158:204–210.

James, A. C., R. B. R. Azevedo, and L. Partridge. 1995. Cellular basis and developmental timing in a size cline of *Drosophila melanogaster*. Genetics 140:659–666.

———. 1997. Genetic and environmental responses to temperature of *Drosophila melanogaster* from a latitudinal cline. Genetics 146:881–890.

Janzen, D. H. 1967. Why mountain passes are higher in tropics. Am. Nat. 101:233–249.

de Jong, G. 1995. Phenotypic plasticity as a product of selection in a variable environment. Am. Nat. 145:493–512.

de Jong, G., and S. Gavrilets. 2000. Maintenance of genetic variation in phenotypic plasticity: the role of environmental variation. Genet. Res. Camb. 76:295–304.

Kearsey, M. J., and H. S. Pooni. 1996. The genetical analysis of quantitative traits. Chapman and Hall, London.

Kennington, W. J., A. S. Gilchrist, D. B. Goldstein, and L. Partridge. 2001. The genetic bases of divergence in desiccation and starvation resistance among tropical and temperate populations of *Drosophila melanogaster*. Heredity 87:363–372.

Kruuk, L. E. B., J. Slate, and A. J. Wilson. 2008. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. Annu. Rev. Ecol. Evol. Syst. 39:525–248.

Lacaze, X., P. M. Hayes, and A. Korol. 2009. Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. Heredity 102:163–173.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evol. Biol. 22:1435–1446.

Lee, C., and Y. Kim. 2009. Estimation of interaction effects among nucleotide sequence variants in animal genomes. Asian Australas. J. Anim. Sci. 22:124–130.

Li, Y., R. Y. Cheng, K. A. Spokas, A. A. Palmer, and J. O. Borevitz. 2014. Genetic variation for life history sensitivity to seasonal warming in *Arabidopsis thaliana*. Genetics 196:569–577.

Li, Y., Y. Huang, J. Bergelson, M. Nordborg, and J. O. Borevitz. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 107:21199–21204.

Liefting, M., A. A. Hoffmann, and J. Ellers. 2009. Plasticity versus environmental canalization: population differences in thermal responses along a latitudinal gradient in *Drosophila serrata*. Evolution 63:1954–1963.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Mather, K., and J. L. Jinks. 1982. Biometrical genetics: the study of continuous variation. Cambridge Univ. Press, Cambridge, U.K.

Maynard Smith, E. 1978. The evolution of sex. Cambridge Univ. Press, Cambridge, U.K.

Mayr, E. 1954. Change of genetic enviroment and evolution. Pp. 157–180 *in* J. Huxley, A. C. Hardy, and E. B. Ford, eds. Evolution as a process. Allen and Unwin, London.

Mendez-Vigo, B., M. Savic, I. Ausin, M. Ramiro, B. Martin, F. X. Pico, and C. Alonso-Blanco. 2016. Environmental and genetic interactions reveal FLOWERING LOCUS C as a modulator of the natural variation for the plasticity of flowering in Arabidopsis. Plant Cell Environ. 39:272–294.

Meiri, S., and T. Dayan. 2003. On the validity of Bergmann's rule. J. Biogeogr. 30:331–351.

Murren, C. J., H. J. Maclean, S. E. Diamond, U. K. Steiner, M. A. Heskel, C. A. Handelsman, C. K. Ghalambor, J. R. Auld, H. S. Callahan, D. W.

Pfennig, et al. 2014. Evolutionary change in continuous reaction norms. Am. Nat. 183:453–467.

Orr, H. A. 1995. The population-genetics of speciation—the evolution of hybrid incompatibilities. Genetics 139:1805–1813.

Overgaard, J., M. R. Kearney, and A. A. Hoffmann. 2014. Sensitivity to thermal extremes in Australian *Drosophila* implies similar impacts of climate change on the distribution of widespread and tropical species. Glob. Change Biol. 20:1738–1750.

Paixao, T., and N. H. Barton. 2016. The effect of gene interactions on the long- term response to selection. Proc. Natl. Acad. Sci. USA 113:4422–4427.

Perkins, J. M., and J. L. Jinks. 1973. Assessment and specificity of environmental and genotype-environmental components of variability. Heredity 30:111–126.

Pooni, H. S., D. T. Coombs, P. S. Virk, and J. L. Jinks. 1987. Detection of epistasis and linkage of interacting genes in the presence of reciprocal differences. Heredity 58:257–266.

R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Remold, S. K., and R. E. Lenski. 2004. Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. Nat. Genet. 36:423–426.

Rice, S. H. 1998. The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. Evolution 52:647–656.

Roff, D. A., and K. Emerson. 2006. Epistasis and dominance: evidence for differential effects in life-history versus morphological traits. Evolution 60:1981–1990.

Santos, J., M. Pascual, P. Simoes, I. Fragata, M. Lima, B. Kellen, M. Santos, A. Marques, M. R. Rose, and M. Matos. 2012. From nature to the laboratory: the impact of founder effects on adaptation. J. Evol. Biol. 25:2607–2622.

Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Syst. 24:35–68.

Scheiner, S. M., and R. F. Lyman. 1989. The genetics of phenotypic plasticity .1. Heritability. J. Evol. Biol. 2:95–107.

Schiffer, M., A. S. Gilchrist, and A. A. Hoffmann. 2006. The contrasting genetic architecture of wing size, viability, and development time in a rainforest species and its more widely distributed relative. Evolution 60:106–114.

Schlichting, C. D., and M. Pigliucci. 1993. Control of phenotypic plasticity via regulatory genes. Am. Nat. 142:366–370.

Sgro, C. M., J. Overgaard, T. N. Kristensen, K. A. Mitchell, F. E. Cockerell, and A. A. Hoffmann. 2010. A comprehensive assessment of geographic variation in heat tolerance and hardening capacity in populations of *Drosophila melanogaster* from eastern Australia. J. Evol. Biol. 23:2484–2493.

Sultan, S. E., and H. G. Spencer. 2002. Metapopulation structure favors plasticity over local adaptation. Am. Nat. 160:271–283.

Sunday, J. M., A. E. Bates, and N. K. Dulvy. 2011. Global analysis of thermal tolerance and latitude in ectotherms. Proc. R. Soc. B. 278:1823–1830.

Ungerer, M. C., S. S. Halldorsdottir, M. A. Purugganan, and T. F. C. Mackay. 2003. Genotype-environment interactions at quantitative trait loci affecting inflorescence development in *Arabidopsis thaliana*. Genetics 165:353–365.

van Heerwaarden, B., and C. M. Sgro. 2011. The effect of developmental temperature on the genetic architecture underlying size and thermal clines in *Drosophila melanogaster* and *D. simulans* from the east coast of Australia. Evolution 65:1048–1067.

Via, S. 1993. Adaptive phenotypic plasticity—target or by-product of selection in a variable environment. Am. Nat. 142:352–365.

Via, S. and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39:505–522.

Via, S., R. Gomulkiewicz, G. de Jong, S. M. Scheiner, C. D. Schlichting, and P. H. Vantienderen. 1995. Adaptive phenotypic plasticity—consensus and controversy. Trends Ecol. Evol. 10:212–217.

Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. Nature 150:563–565.

Wade, M. J., and C. J. Goodnight. 1998. Perspective: the theories of Fisher and Wright in the context of metapopulations—when nature does many small experiments. Evolution 52:1537–1553.

Wagner, G. P., G. Booth, and H. C. Bagheri. 1997. A population genetic theory of canalization. Evolution 51:329–347.

Weber, S. L., and S. M. Scheiner. 1992. The genetics of phenotypic plasticity .4. Chromosomal localization. J. Evol. Biol. 5:109–120.

Westerman, J. M. 1971a. Genotype-environment interaction and developmental regulation in *Arabidopsis thaliana* .4. Wild material—analysis. Heredity 26:383–395.

———. 1971b. Genotype-environment interaction and developmental regulation in *Arabidopsis thalinana* .2. Inbred lines—analysis. Heredity 26:93–106.

Whitlock, M. C., P. C. Phillips, F. B. G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks and epistasis. Annu. Rev. Ecol. Syst. 26:601–629.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:0097–0159.

Wu, R. L. 1998. The detection of plasticity genes in heterogeneous environments. Evolution 52:967–977.

Zhou, S. S., T. G. Campbell, E. A. Stone, T. F. C. Mackay, and R. R. H. Anholt. 2012. Phenotypic plasticity of the *Drosophila* transcriptome. PLoS Genet. 8:e1002593.

Associate Editor: L.M. Chevin
Handling Editor: P. Tiffin

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Result for *t* tests and one-way analysis of variance examining divergence between the parental lines (*t* tests) and across the cohorts (ANOVA) in elevation, slope, and curvature for heat resistance, cold resistance, wing centroid size, and thorax size between tropical and temperate populations of *Drosophila melanogaster* and *D. simulans.*

Table S2. Information on the number of models used to construct the 95% confidence set of models that contain the minimum number of models whose Akaike weights (*wi*) sum to 0.95 and the *wi* of the best model.

Table S3. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature against alternative models including the CGEs identified as important for the other reaction norm parameters for heat and cold resistance in *Drosophila melanogaster* and *D. simulans*.

Table S4. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature against alternative models including the CGEs identified as important for the other reaction norm parameters for wing centroid size in female and male *Drosophila melanogaster* and *D. simulans.*

Table S5. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature against alternative models including the CGEs identified as important for the other reaction norm parameters for thorax size in female and male *Drosophila melanogaster* and *D. simulans.*

Table S6. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature in females against alternative models including the CGEs identified as important for the same reaction norm parameter in males (and vice versa) for wing size in *Drosophila melanogaster* and *D. simulans.*

Table S7. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature in females against alternative models including the CGEs identified as important for the same reaction norm parameter in males (and vice versa) for thorax size in *Drosophila melanogaster* and *D. simulans.*

Table S8. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature in *Drosophila melanogaster* against alternative models including the CGEs identified as important for the same reaction norm parameter in *Drosophila simulans* (and vice versa) for heat.

Table S9. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature in *Drosophila melanogaster* against alternative models including the CGEs identified as important for the same reaction norm parameter in *Drosophila simulans* (and vice versa) for wing centroid size.

Table S10. AICc scores for models comparing the fit of composite genetic effects (CGEs) identified as important for divergence in elevation, slope, and curvature in *Drosophila melanogaster* against alternative models including the CGEs identified as important for the same reaction norm parameter in *Drosophila simulans* (and vice versa) for thorax size.

Figure S1. Mean elevation, slope, and curvature for heat knockdown time in *Drosophila melanogaster* (a, b, c) and *Drosophila simulans* (d, e, f), and cold in *D. melanogaster* (g, h, i) as a function of the proportion of genes derived from P1, the temperate parent (Melbourne) in each cross.

Figure S2. Mean elevation, slope, and curvature for wing centroid size in *Drosophila melanogaster* females (a, b, c) and males (d, e, f), and in *Drosophila simulans* females (g, h, i) and males (j, k, l) as a function of the proportion of genes derived from P1, the temperate parent (Melbourne) in each cross.

Figure S3. Mean elevation, slope, and curvature for thorax size in *Drosophila melanogaster* females (a, b, c) and males (d, e, f), and in *Drosophilasimulans* females (g, h, i) and males (j, k, l) as a function of the proportion of genes derived from P1, the temperate parent (Melbourne) in each cross.