The predictive potential of key adaptation parameters and proxy fitness traits between benign and stressful thermal environments

Jennifer M. Cocciardi[1*], Eleanor K. O'Brien[2,4], Conrad J. Hoskin[1], Henry Stoetzel[3], Megan Higgie[1]

[1]College of Science and Engineering, James Cook University, Douglas, QLD 4811, Australia, [2]School of Biological Sciences, University of Bristol, Bristol BS8 1TL, UK, [3]School of Earth and Environmental Sciences, University of Queensland, St Lucia, QLD, 4072, Australia, [4]Current address: Centre for Precision Health, School of Medical & Health Sciences, Edith Cowan University, Joondalup, WA 6027, Australia.

*Corresponding author details: Jennifer M. Cocciardi, College of Science and Engineering, James Cook University, Douglas, QLD 4811, jenny.cocciardi@my.jcu.edu.au

Running headline: Genetic variances and covariances across environments

Word count: 6998

Abstract

Understanding the adaptive potential of a species is key when predicting whether a species can contend with climate change. Adaptive capacity depends on the amount of genetic variation within a population for relevant traits. However, genetic variation changes in different environments, making it difficult to predict whether a trait will respond to selection when not measured directly in that environment. Here, we investigated how genetic variances, and phenotypic and genetic covariances, between a fitness trait and morphological traits changed between thermal environments in two closely-related *Drosophila*. If morphological traits strongly correlate with fitness, they may provide an easy-to-measure proxy of fitness to aid in understanding adaptation potential. We used a parent-offspring quantitative genetic design to test the effect of a benign (23°C) and stressful (28°C) thermal environment on genetic variances of fecundity and wing size and shape, and their phenotypic and genetic covariances. We found genetic variances were higher within the stressful environment for fecundity but lower within the stressful environment for wing size. We did not find evidence for significant phenotypic correlations. Phenotypic and genetic correlations did not reveal a consistent pattern between thermal environments *or* within or between species. This corroborates previous research and reiterates that conclusions drawn in one environment about the adaptive potential of a trait, and the relationship of that trait with fitness, cannot be extrapolated to other environments *or* within or between closely-related species. This confirms that researchers should use caution when generalising findings across environments in terms of genetic variation and adaptive potential.

Introduction

Climate change is causing increased temperatures that will impose stress on species (Thomas et al. 2004). Many species lack the ability to disperse to more optimal environments (Bellard et al. 2012; Ceballos et al. 2017), and will have to adapt to the stressful temperatures to survive in the long-term (Thomas et al. 2004; Hoffmann and Sgrò 2011). Adaptation potential will depend on the amount of genetic variation in traits relevant to the selection imposed by environmental change (Fisher 1930; Falconer and Mackay 1996), and adaptation will need to be rapid given the speed of human-induced climate change. Understanding the adaptive potential of species, especially those currently living close to their upper thermal limits, is therefore crucial in today's changing climate (Urban et al. 2016; Funk et al. 2019; Shaw 2019).

Importantly, genetic variation is context-dependent — meaning the amount of genetic variation in a trait in a given population can change under different environments (Falconer and Mackay 1996; Hoffmann and Schiffer 1998; Sgrò and Hoffmann 1998a, c; Hoffmann and Merilä 1999). Short-term environmental changes can play an important role in adaptive evolution (Wood and Brodie 2016) and can induce a similar or larger change in genetic variance than changes to the genetic architecture that accumulate over hundreds of generations between populations (for review, see Wood and Brodie 2015). Increases in environmental variability, such as those predicted with climate change, will therefore directly affect the rate of evolution of a trait — as environments get warmer, not only may the type of selective pressure change, but also the potential for the trait to respond to selection. This is important because as researchers aim to determine whether species can adapt to climate change, the changing climate itself may increase or decrease adaptation potential.

Much research has focused on examining whether there is a consistent pattern to changes in expression of genetic variance (for reviews, see Sgrò and Hoffmann 2004; Rowiński and Rogell 2017; Fischer et al. 2020). However, there is no consensus on whether stressful conditions increase or decrease the expression of genetic variance. The majority of studies focus on quantifying genetic variance by calculating heritability (h^2), which describes the relative amount of genetic variance due to additive effects standardised by total phenotypic variance. Heritability can then be used to predict the magnitude of the response to selection via the breeder's equation. These studies show both increases (e.g., Sgrò and Hoffmann 1998a, c; Swindell and Bouzat 2006) and decreases (e.g., Hoffmann and Schiffer 1998; Bubliy et al. 2001; Kristensen et al. 2015) in heritability under stressful conditions. Increased heritability may result from novel genetic variance that is expressed when exposed to new conditions (i.e., 'cryptic genetic variance'; see review by Hoffmann and Schiffer 1998; but also see Swindell and Bouzat 2006). Decreased heritability may result from low cross-environment genetic covariances (Fischer et al. 2020), or environmental variance increasing while other variance components remain the same (for example see Hoffmann and Schiffer 1998). Recently, studies have recommended quantifying genetic variance using parameters standardized by the trait mean — such as coefficient of additive variance (CV_A) and its square, evolvability (I_A) — because estimates of heritability, which are standardised by the phenotypic variance, can be influenced by sources of non-genetic environmental variation that may preclude comparison across environments and traits (Houle 1992).

Assessing the effect of a changing environment on genetic variance is further complicated when attempting to measure genetic variance across different environments for fitness. Direct fitness (reproductive success) is often difficult to measure in the wild because of uncontrolled and unmeasured factors (Orr 2009), and in the laboratory due to time and logistical constraints (Rosenberg 1982; Nguyen and Moehring 2015). Instead, a

morphological trait that strongly correlates with fitness, and is more easily measured, may provide a good proxy when fitness measures are difficult to obtain. If phenotypic correlations between a morphological trait and fitness are strong, researchers can use the easier-to-measure trait to predict genetic variation of fitness across different environments (Arnold 1983).

More importantly, a strong phenotypic correlation may indicate that two traits are genetically linked through physical linkage, pleiotropy, or linkage disequilibrium (Cheverud 1988; Conner and Via 1992; Roff 1995; Blows and Hoffmann 2005). This means a positive genetic correlation between traits could aid adaptation to novel environments if selection favours that trait combination through augmenting the effect of selection on the correlated fitness trait (Blows and Hoffmann 2005; Agrawal and Stinchcombe 2009; Walsh and Blows 2009; Holman and Jacomb 2017). Therefore, determining genetic correlations of traits with fitness is an important part of the puzzle when predicting evolutionary potential.

However, much like genetic variation in individual traits, phenotypic and genetic covariances between traits (or between a trait and fitness) can vary depending upon the environment in which they are measured (Sgrò and Hoffmann 2004) — meaning measurements obtained in one environment cannot necessarily be generalized to other environments. For example, in a beetle, adult female body mass and her egg size were positively correlated on one host-plant species and negatively correlated on a different host-plant species (Czesak and Fox 2003). Changes to genetic correlations can result within novel environments due to genotype-environment interactions — where genes that affect a trait in one environment may not be influential in a different environment (Sgrò and Hoffmann 2004). In some instances, the loci that contribute to covariances through pleiotropy or physical linkage have specifically been found to be influenced by environmental effects (e.g., Hausmann et al. 2005; Gutteling et al. 2007). However, more empirical data are needed to

115 understand whether there are patterns to how genetic variances and covariances of
116 morphological and fitness traits vary across thermal environments (Rowiński and Rogell
117 2017; Fischer et al. 2020).

118 *Drosophila* are often used to investigate genetic variances due to their short
119 generation time and ability to produce large numbers of offspring that allow for quantitative
120 genetic experimental designs. Fecundity is a commonly assessed fitness trait in *Drosophila*.
121 However, measuring fecundity can often prove time- and labour-intensive and logistically
122 challenging. Ecological theory assumes that body size is correlated with fecundity, with
123 larger individuals exhibiting a higher fecundity (Chiang and Hodson 1950; Santos et al. 1992;
124 Robertson 1956), and wing length has been shown to phenotypically correlate with fecundity
125 (Tantawy and Vetukhiv 1960; Woods et al. 2002). However, two key studies examining the
126 relationship of wing length and fecundity in *Drosophila* when exposed to stressful
127 environments found mixed evidence. Sgró and Hoffmann (1998a) did not detect a significant
128 positive phenotypic or genetic correlation in a cold-stress, heat-stress, or benign environment.
129 They also did not find a significant genetic cross-environment correlation (parents raised in
130 one environment and offspring raised in a different environment) between cold-stress, heat-
131 stress, or benign environments (Sgrò and Hoffmann 1998a) — meaning that they did not find
132 a correlation between wing length and fecundity among and between any experimental
133 environment. Conversely, Woods et al. (2002) found significant positive phenotypic
134 correlations (for two of three generations) and significant positive genetic correlations
135 between wing length and fecundity in a stressful environment, but not in a benign
136 environment.

137 With advances in technology over the past decade (i.e., advances in microscopic
138 imaging and digitizing), more intricate morphological traits such as wing size and wing shape
139 have been increasingly used in place of wing length. However, very few studies have

140 examined genetic variation and heritability in wing size and shape (Gilchrist and Partridge
141 1999; Hoffmann and Shirriffs 2002; Moraes et al. 2004); and, to our knowledge, only one has
142 examined the phenotypic and genetic correlations of wing size with fecundity (e.g., Woods et
143 al. 2002). Wing size and wing shape in *Drosophila* have a polygenic basis independent of
144 one another (Carreira et al. 2011), so phenotypic and genetic correlations of each of these
145 traits with fecundity may differ. Wing size exhibits a history of directional selection in
146 *Drosophila*, whereas wing shape has been shown to undergo optimizing selection (Gilchrist
147 and Partridge 2001). Although most of the fundamental research uses wing length as a trait
148 that is highly correlated to thorax size (and therefore body size; Chiang and Hodson 1950;
149 Tantawy and Vetukhiv 1960; Santos et al. 1992; Woods et al. 2002), wing size may be a
150 better indicator of overall body size because it is a product of more complex interactions
151 between the different wing compartments (i.e., anterior and posterior compartments; Guerra
152 et al. 1997; Gilchrist and Partridge 1999). Hence, wing size may account for a greater
153 proportion of variation than wing length alone. Wing shape is important for flight
154 performance in *Drosophila* and has been shown to exhibit high heritability (Hoffmann and
155 Shirriffs 2002; Moraes et al. 2004). However, whether selection occurs on wing shape itself,
156 or whether wing shape is correlated with another trait under selection is unknown (Gilchrist
157 and Partridge 2001).

158 Temperature as a stressor is contextually important in today's climate, but it has only
159 been used in one *Drosophila* study to assess whether genetic correlations exist between
160 fecundity and wing length, and whether heritability changes between different thermal
161 regimes (i.e., *D. melanogaster*; Sgrò and Hoffmann 1998a with the same data used in
162 Woods et al. 2002). Here, we focused on whether genetic variances in fecundity change
163 across thermal environments, and whether a morphological trait that may be a good proxy of
164 fitness in one environment was also a good proxy in a stressful thermal-environment. We

examined the consistency of heritability, coefficient of additive genetic variance, and evolvability between thermal environments (one benign and one stressful), generations, and within and between two sibling species of *Drosophila*. A strength of this study is that we assessed both life history and morphology traits in two closely-related species to see whether this pattern was conserved. We also assessed the phenotypic and genetic covariances of these traits. The correlation of body morphology with fitness informs us about the strength and direction of selection. This is important because patterns of selection in one environment may not reflect similar responses in another environment.

Methods

Experimental populations

Two sibling species of fruit fly found along the east coast of Australia were used in this study: *Drosophila serrata*, a generalist species found in forested areas; and *D. birchii*, a specialist species confined to tropical rainforest ecosystems (Schiffer and Mcevey 2006; Higgie and Blows 2008). Mass bred populations from two different geographical areas for each species were used. Each mass bred population was originally created by breeding the offspring of ten isofemale lines collected from field sites within Queensland, Australia. *Drosophila birchii* flies were collected from Paluma National Park (19° 0'16.27"S, 146°12'35.59"E) and Mt. Lewis National Park (16°35'30.36"S, 145°16'27.78"E). *Drosophila serrata* flies were collected from Paluma National Park (19° 0'16.27"S, 146°12'35.59"E) and Raglan Creek (23°42'49.74"S, 150°49'0.10"E). All flies were collected between February and May 2016. Isofemale lines were maintained in controlled laboratory conditions for 18 generations before mass bred populations were created. All stocks were maintained at large

190 population sizes ($N > 1000$) to retain natural genetic variation. Flies were reared on standard
191 *Drosophila* food that contained sugar, yeast, and agar as described in Higgie and Blows
192 (2008). All flies were reared under constantly controlled laboratory conditions of 23°C ±
193 1°C, 50% relative humidity (RH), and 12 h light:dark cycles.
194
195 *Quantitative genetic experimental design*
196
197 A parent-offspring breeding design was used to assess heritability and phenotypic and genetic
198 covariances of fecundity and wing morphology at a benign (23°C) and a stressful (28°C)
199 temperature (Fig. 1). The benign temperature (23°C) represents an approximate average
200 temperature each species experiences across their range both temporally and spatially, as well
201 as the optimal rearing temperature in the laboratory. A temperature of 28°C was chosen as a
202 stressful thermal environment as it was found to be within the upper margin of the thermal
203 niche for *D. birchii* and to place stress upon *D. serrata* (from pilot studies showing reduced
204 survival). As such, full development was expected in both species at this temperature.
205 Two generations before the start of the experiment, density-controlled mass bred
206 populations were created for each species and population by sexing 25 virgin females and 25
207 virgin males from the laboratory stock and placing them in one 300 mL bottle with 100 mL
208 of food. This was repeated three times for each species and population. Flies were removed
209 from each replicate bottle after 72 h and bottles were carded for pupation. Offspring were
210 collected at random and sexed to subsequently create family lines and stock mass bred
211 populations for each species and treatment.
212 One generation before the start of the experiment (i.e., P generation; Fig. 1), virgin
213 offspring were sexed from the density-controlled mass bred populations using CO_2
214 anesthetization. Flies were placed in 100 mL holding vials with 5 mL of food for 72 h to

allow for sexual maturation and full recovery from anesthetization, with 5 individuals per holding vial. After this, one male and one female were randomly collected and placed in a 100 mL glass vial with 10 mL of food, stoppered with a porous stopper, and directly placed in an incubator set to the relevant temperature for each thermal environment treatment. Humidity inside the vials with the stoppers on remained at approximately 90% RH, and a 12 h light:dark cycle was maintained. This was carried out for 50 family replicates for each species, population, and treatment. Mating pairs were allowed to mate for 48 hours before being removed from the vial. This ensured all experimental flies were reared in a controlled and low-density environment. In addition, three low-density stock bottles containing 10 females and 10 males were created and maintained for both the parent and offspring generations to provide a supply of males for mating to assess fecundity (i.e., male mass breds; Fig. 1). Three stock mass bred bottles were maintained in each thermal environment and males were randomly collected from each bottle and mated with a female from the same experimental rearing temperature.

Fecundity measurements

Virgin female offspring of each family replicate vial were sexed under light anesthetization and placed in holding vials for 72 h. One female (i.e., dam) from each F_1 family was randomly selected and placed in an empty vial with one virgin male collected from the male stock bottles. Each vial contained a small spoon with 2 mL of food to provide a medium for oviposition. The food was dyed green to aid in counting eggs, and a drop of a live yeast-water solution (1 g baker's yeast:10 mL water) was spread over it to promote ovipositing. Vials were immediately placed within their temperature treatment and flies were allowed to mate for 24 h. After 24 h, the spoon was removed and immediately frozen at -19°C for eggs

to be counted at a later time, and replaced with a new spoon. This was repeated every 24 h for three days and a cumulative fecundity count was obtained. Cumulative fecundity measurements from the first three days of maturity are significantly correlated with lifetime reproductive success of female *Drosophila* (Pekkala et al. 2011; Nguyen and Moehring 2015). After 72 h, the mating pair was transferred to a rearing vial with 10 mL of food and allowed to mate for the next 48 h period before being removed. Females were then immediately frozen for wing morphology measurements. Daughters of these pairs were collected from each vial and one virgin female offspring from each mating pair was assessed for fecundity and wing traits using the same methods described above.

Fecundity was scored using a microscope and click counter by counting the number of eggs on each spoon. Approximately one quarter of spoons were counted twice, at random, to assess repeatability; a positive correlation close to 1 indicated that counting was highly repeatable between measurements ($r = 0.994$; $P < 0.001$; $N = 81$).

Wing morphometrics

All dams and daughters were frozen at nine days old to assess wing size and shape. Left wings were removed using fine forceps and mounted on microscope slides with double-sided tape. Wings were photographed using a Leica Image microscope and Leica Application Suite software (LAS v. 3.8). Images were randomized and collated as a TPS file using *tpsUtil* (Rohlf 2010a). Landmarks were placed on ten consistent morphometric wing features of each image (Supplementary Fig. 1) using the program *tpsDig2* (Rohlf 2016). Outliers and landmarking errors were identified using *tpsRelW* (Rohlf 2010b) and corrected or removed before wing measurements were computed.

Landmarked coordinates underwent a Generalized Procrustes Analysis (GPA) superimposition (Rohlf and Slice 1990), where wing size and alignment are adjusted for by superimposing images upon one another over an average configuration. The GPA superimposition has been found to produce estimates with the least amount of error in a study on geomorphometrics (Rohlf 2003). The square root of the summed squared distance between centroid configuration and landmarks is known as the centroid size and provides a measure of overall size (Rohlf and Slice 1990; Rohlf 2000). Although size effects should be removed via the GPA, a correlation between shape and size might still occur (known as allometry), and hence this was also assessed.

In addition to centroid size, the GPA computes a set of Procrustes residuals for each landmark. A principle component analysis (PCA) was conducted upon these to identify variation components, which can be used to describe a single axis of variation in wing shape among individuals (Adams et al. 2004; Zelditch et al. 2004; Gómez et al. 2009). In this instance, the PCA is equivalent to a relative warp analysis because the variation between landmarks was not weighted by bending energy (Zelditch et al. 2004), so PCA scores are equivalent to relative warp (RW) scores. As per common practice, RW scores that explained greater than 5% of variation were used as shape variables (Zelditch et al. 2004; Gómez et al. 2009) to analyse differences in wing shape using the *geomorph* package (Adams et al. 2020) in R.

Analysis

Data was checked for outliers, homogeneity, normality, and independence as outlined in the protocol described in Zuur et al. (2010) and analyses were performed using the statistical program R (R Core Team 2019). Mean trait values and phenotypic variances (calculated as

289 squared standard errors of the mean trait values) were calculated for fecundity, wing size, and
290 wing shape for both the dam and daughter generations. To test whether thermal environment
291 had an effect on mean trait values and on phenotypic variances, a two-way ANOVA (for
292 thermal environment and generation and its interaction) or a generalized least square model
293 was conducted for each trait and metric, depending on data structure. Population and its
294 interaction with thermal environment was also included when significant. For the
295 multivariate measure of shape (RW score matrix), a permutational MANOVA (also known as
296 a Procrustes ANOVA; Goodall 1991) was used to test for differences in wing shape between
297 thermal treatments, populations, and generations. All analyses were conducted separately for
298 each species.

299 Narrow sense heritability (h^2) for each trait was calculated from a regression of
300 offspring trait values on maternal trait values (Falconer and Mackay 1996). As we conducted
301 only a single-parent regression, the phenotypic resemblance is equal to half of the genetic
302 variation and thus the slope parameter estimate β represents:

303 $$\beta = 1/2(\frac{V_A}{V_P})$$

304 $$\beta = 1/2h^2$$

305 and so, heritability is equal to twice the slope of the regression line (Falconer and Mackay
306 1996). In parent-offspring regression, estimates can be greatly skewed if variances found in
307 the parental generation and offspring generation differ (Falconer and Mackay 1996). To
308 overcome this, we standardized all traits to a mean of zero and standard deviation of one
309 prior to computation of heritability and genetic covariances (Sgrò and Hoffmann 1998b). The
310 significance of deviations of heritability estimates from zero were assessed using an *F*-test
311 and all *P*-values were adjusted for using the False Discovery Rate method (Benjamini and
312 Hochberg 1995). Standard errors of heritability were obtained directly from the regression
313 model.

To obtain an overall estimate of heritability for wing shape, we followed the equations set forth in Monterio et al. (2002) for estimating heritability from a parent-offspring regression on a multivariate trait (i.e., wing shape with all RW scores included). This was done by first obtaining the coefficient of determination (R^2) from a multivariate linear regression of offspring RW scores onto dam RW scores. The square root of the coefficient of determination (R) was then used in the following formula:

$$\beta = R\frac{S_O}{S_P}$$

where β is the multivariate regression coefficient, S_O is the standard deviation of the offspring trait, and S_P is the standard deviation of the parental trait. The multivariate regression coefficient multiplied by two is then equal to the heritability of the trait (to account for only having half of the genetic variation due to the single-parent-offspring comparison). Significance of deviation from zero for multivariate shape heritability was assessed using a Wilks' lambda test (Zelditch et al. 2004; Gómez et al. 2009).

In addition, coefficients of genetic variation (CV_A) and evolvabilities (I_A) were calculated following Houle (1992) as:

$$CV_A = \frac{\sqrt{V_A}}{\bar{X}}$$

$$I_A = \frac{V_A}{\bar{X}^2}$$

Because we did not directly calculate V_A in this analysis, we obtained estimates based on the method of Garcia-Gonzalez et al. (2012). V_A estimates were calculated by multiplying the total phenotypic variance (V_P) of each trait mean by the narrow-sense heritability (h^2), since $V_A = h^2 \times V_P$ (Falconer and Mackay 1996). This is an alternative way to calculate CV_A when researchers do not have the sire variance component (V_{sire}) or another direct measure of V_A (Garcia-Gonzalez et al. 2012). Standardized data cannot be used to calculate CV_A and I_A

because a scaling correction to a zero mean produces a meaningless comparison and undefined value when dividing by the trait mean a second time (Garcia-Gonzalez et al. 2012). The above methods were therefore only performed on non-standardized data and CV_A and I_A values were not calculated for RW scores of wing shape as these are standardized.

The phenotypic correlation among each pair of traits was calculated as the Pearson correlation coefficient. Genetic covariances (cov_{XY}) were obtained by regressing one trait in the parental generation onto the other trait in the offspring generation, in both directions (Supplementary Table 5), adjusting for relationship, and taking the mean of the adjusted Pearson correlation coefficients as suggested by Falconer and Mackay (1996). Genetic correlations were then calculated using the genetic covariances and the following equation:

$$r_G = \frac{cov_{XY}}{\sqrt{cov_{XX} cov_{YY}}}$$

where cov_{XY} is the genetic 'cross-covariance' and cov_{XX} and cov_{YY} are the parent-offspring covariances for the individual traits. Standard errors for genetic correlations were calculated using an approximate formula as proposed by Reeve (1955), Robertson (1959) and explained in Falconer and Mackay (1996):

$$\sigma_{r_G} = \frac{1 - r_G^2}{\sqrt{2}} \sqrt{\left[\frac{\sigma_{(h_X^2)}\, \sigma_{(h_Y^2)}}{h_X^2 h_Y^2}\right]}$$

All correlations were estimated using linear regression models that initially included the main effects of temperature and population and an interaction between them, with interaction and population terms removed if they were non-significant. In the majority of cases, population was not significant and this allowed for one correlation value per species.

Results

Mean trait values differed significantly between thermal environments for each species and generation ($p = < 0.001$). Rearing in a stressful thermal environment resulted in lower fecundity (Supplementary Table 1 and Supplementary Fig. 2), smaller wing size (Supplementary Table 2 and Supplementary Fig. 3), and a rounder, less elongated wing shape when adjusted for size (Supplementary Table 3 and Supplementary Fig. 4) across all species, populations, and generations.

How does genetic variation change in a stressful thermal environment?

Fecundity

Phenotypic variation in fecundity did not differ significantly between thermal environments, but was slightly higher within the stressful environment than the benign environment (Table 1). CV_A estimates could not be calculated for *D. birchii* within the stressful temperature because unexpectedly the offspring did not emerge in this treatment. CV_A and evolvability (I_A) estimates for fecundity were higher than for morphological traits in all instances, and slightly higher in the stressful environment than in the benign environment in *D. serrata* (Table 1 and Fig. 2). Fecundity was found to have a low heritability overall that was slightly higher under the benign than stressful thermal environment in *D. serrata* (Table 1 and Fig. 2).

Morphological wing traits

Phenotypic variation in wing size differed significantly between thermal environments for dams of both species (*D. serrata*: $P = 0.02$; *D. birchii*: $P = 0.005$), and was slightly higher within the benign environment (Supplementary Table 2). Heritability, evolvability, and CV_A estimates were higher within the benign environment than the stressful environment in *D. serrata*, and heritability values were overall much higher for wing size compared to fecundity (Table 1 and Fig. 2).

Phenotypic variation in wing shape variables significantly differed between thermal environments for all RW scores in *D. serrata* ($P < 0.005$), but did not differ between thermal environments in *D. birchii*. The direction and magnitude of changes in phenotypic variances did not show a consistent pattern across thermal environments (Supplementary Table 3). Wing shape heritability increased within the stressful environment in *D. serrata* (Table 1 and Fig. 2). Heritability in all instances was much higher than for fecundity. In addition, wing size and wing shape evolvabilities and CV_A estimates were all very low compared to fecundity (Fig. 2).

How does phenotypic correlation of traits change in a stressful environment?

There were no significant phenotypic correlations found between fecundity and wing size after correction for multiple comparisons (Table 2 and Fig. 3).

Fecundity and wing shape traits exhibited mixed and inconsistent results (Fig. 3 and Supplementary Table 4). There was only one significant phenotypic correlation found between fecundity and a wing shape variable within the daughter generation of the Mt. Lewis population of *D. serrata* under a stressful thermal environment. Allometry was found within

406 the benign environment for the daughter generation of *D. serrata*, indicating wing size and
407 wing shape in this instance are still slightly correlated even after removing effects of size
408 during the GPA analysis (Supplementary Table 4).
409
410 *How do genetic covariances and correlations change with environmental stress?*
411
412 There was no consistent trend detected for genetic covariances and no consistent pattern in
413 genetic correlations (i.e., genetic covariances standardized by individual trait covariances)
414 between species and thermal environments (Fig. 3; Supplementary Tables 5, 6). Genetic
415 correlations in *D. birchii* were generally low ($-0.32 < r_G < 0.46$), while *D. serrata* traits
416 exhibited high positive and negative genetic correlations, but this was not consistent across
417 environments (Fig. 3).
418 We did find highly negative and highly positive genetic correlations between
419 fecundity and wing morphometries in *D. serrata* (including values of ± 1.00). However, these
420 often had very wide standard errors and were not always significant. In the benign
421 environment for *D. birchii*, we found a significant positive genetic correlation between wing
422 size and a wing shape variable (RWb-2; $r_G = 0.46 \pm 0.14$ SE; $P < 0.01$). In the benign
423 environment for *D. serrata*, we found a significant negative genetic correlation between
424 fecundity and wing size ($r_G = -1.00 \pm 0.08$ SE; $P < 0.001$) and a significant positive (RWs-1;
425 $r_G = 0.75 \pm 0.16$ SE; $P < 0.0001$) and negative correlation between fecundity and a wing
426 shape variable (RWs-4; $r_G = -0.92 \pm 0.10$ SE; $P = 0.0001$). In the stressful environment for *D.*
427 *serrata*, we found a significant positive correlation between fecundity and wing size ($r_G =$
428 0.84 ± 0.29 SE; $P < 0.05$) and fecundity and a wing shape variable (RWs-4; $r_G = 1.00 \pm 0.13$
429 SE; $P < 0.0001$; Fig. 3).
430

Discussion

The amount of genetic variation in a trait is important for predicting responses of populations and species to climate change as it determines the extent to which a trait can evolve via selection. However, genetic variance and heritability change between environments (Falconer and Mackay 1996; Hoffmann and Schiffer 1998; Hoffmann and Merilä 1999) potentially due to increased environmental variance and reduced additive genetic variance, genotype-by-environment interactions that affect cross-environment genetic correlations, and cryptic genetic variation (Fischer et al. 2020). It is essential to recognise and incorporate this consideration into climate change adaptation research (Shaw 2019), but consistent and predictable patterns have not been detected. It is unclear whether such patterns exist or whether genetic variance and heritability must always be considered in the context of specific traits, populations and environments. Here, we show that temperature stress can alter the heritability, coefficient of additive genetic variation, and evolvability, of both fecundity and morphological traits in two closely-related species of *Drosophila* (one being a generalist and one being a specialist). However, we found no consistent pattern in the direction of change in additive genetic variance and phenotypic and genetic covariances across thermal environments.

First, we confirmed that the warmer ('stressful') thermal environment did indeed induce stress in both species, as demonstrated by lower fecundity, smaller size, and a significantly different wing shape (Supplementary Tables 1–3 and Supplementary Figs. 2–4). In addition, the specialist species (*D. birchii*) failed to develop offspring within the stressful thermal environment. Although there were no experimental differences between the dam and daughter generation that might have caused this, it is possible that maternal effects induced by development within a stressful environment prevented the production of viable offspring.

456 This could alternatively be a paternal effect as it has been shown that *D. birchii* sperm is very
457 sensitive to thermal stress during development (Saxon et al. 2018), and because we attempted
458 to control for maternal effects by developing the initial parental generation (i.e., P in Fig. 1)
459 in the same thermal environment as dams and daughters. Although not a direct aim of this
460 paper, measuring the viability of offspring within a stressful environment is relevant to many
461 evolutionary studies (both in the laboratory and in the field). This is because many studies
462 estimate fitness by measuring the number of offspring directly, but the viability of those
463 offspring are what will maintain the long-term fitness of a population.

464 Second, we found lower heritability in the stressful compared to the benign thermal
465 environment for fecundity and wing size, although not in wing shape in *D. serrata* (Table 1;
466 Fig. 2B). This corroborates a large number of previous studies that show heritability declined
467 under stressful conditions (e.g., Hoffmann and Schiffer 1998; Kristensen et al. 2015; and
468 reviewed in Hoffmann and Parsons 1991; Hoffmann and Merilä 1999; Charmantier and
469 Garant 2005; Rowiński and Rogell 2017). This has important implications for species living
470 close to their upper thermal limits (like many species in the tropics (Deutsch et al. 2008;
471 Kingsolver et al. 2013) because even a small change in environmental conditions may induce
472 a large amount of stress, and these results suggest adaptive potential is reduced under
473 stressful temperatures.

474 However, heritability has been shown to have inherent issues when comparing
475 between environments, as non-additive genetic and environmental variation contribute to it
476 (Houle 1992). To address this problem, we also investigated the coefficient of additive
477 genetic variation (CV_A) and evolvability (I_A). These are often more appropriate estimates to
478 use when comparing genetic variation and evolvability across traits and environments, as
479 they are not affected by non-additive sources of environmental variance (Houle 1992; Bubliy
480 and Loeschcke 2002; Garcia-Gonzalez et al. 2012). Specifically, while heritability tells us the

481 expected absolute change in a trait mean under a given strength of selection from one
482 generation to the next, evolvability predicts the relative change as a percentage of the trait
483 mean (Hansen et al. 2003; Hansen et al. 2011; Garcia-Gonzalez et al. 2012). CV_A and I_A were
484 higher under the stressful environment for fecundity, while the opposite was true for wing
485 size (Fig. 2D). Therefore, while the heritability values suggest that the response to selection
486 on fecundity and wing size will decrease under stressful temperatures, CV_A and I_A suggest
487 that fecundity has greater relative evolutionary potential under the stressful environment than
488 the benign environment and the opposite is true for wing size (Fig. 2). Although it seems
489 heritability and CV_A values may be contradictory, it could be that while the absolute change
490 in fitness (i.e., heritability) will be less in the stressful environment for fecundity, there will
491 be a greater relative increase in fitness in the stressful environment because mean fitness is
492 lower — but this could result in a smaller absolute change in trait mean thus corroborating
493 the heritability results. However, CV_A and I_A values are still important metrics to consider
494 because they will always change in the same direction as V_A. Alternatively, heritability does
495 not necessarily change in the same direction as V_A because effects that increase V_A often
496 increase total variance, which in turn will decrease heritability.
497 An increase in additive genetic variance under stressful temperatures for the measured
498 fitness trait (i.e., fecundity) is advantageous for these *Drosophila* species, both of which live
499 near critical thermal limits (Kellermann et al. 2009; Overgaard et al., 2011). Interestingly,
500 these results are consistent with a recent meta-analysis (Rowiński and Rogell 2017), which
501 showed that the coefficient of genetic variance (CV_A) was higher under stressful conditions
502 for life history traits but not for morphological traits. In terms of wing size, it should be noted
503 that the measured CV_A differed from values previously measured in *D. birchii* (Kellermann et
504 al. 2006), where CV_A was relatively two-fold higher than what we found here. However, in
505 the previous experiment (Kellermann et al. 2006), wing size was measured from flies reared

506 in a benign environment at 25°C (compared to the benign environment measured in this study
507 at 23°C), potentially indicating that even a slight difference in thermal environments can
508 affect estimates of genetic variance. The reported V_A and V_P values indicate differences in
509 CV_A between this study and Kellermann et al. (2006) are due to an increased V_A in their study
510 and not a difference in trait mean that could also induce larger CV_A values (if the trait mean
511 was lower). Collectively, this, along with the other results discussed here, reveal that
512 environmental interactions (that are included in estimating heritability but not CV_A and I_A),
513 potentially play a very large role in shaping the amount of additive genetic variance that
514 selection can act upon.

515 Overall, our heritability values are similar to those reported for fecundity, wing size,
516 and wing shape for *Drosophila* (Gilchrist and Partridge 1999; Hoffmann and Shirriffs 2002;
517 Moraes et al. 2004; Kellermann et al. 2006). Additionally, we examined the differences in
518 genetic variation between fecundity and morphological traits, since patterns in heritability
519 and additive variance (CV_A and I_A) were contradictory. We found that heritabilities were
520 higher for the wing morphology traits than for fecundity (Fig. 2A, B). This coincides with the
521 majority of literature that show morphological traits often have higher heritabilities than life
522 history traits (Mousseau and Roff 1987; but for opposing example see Sgrò and Hoffmann
523 1998c). In direct contrast to this, CV_A and I_A were both magnitudes larger for fecundity than
524 what was found for wing morphology (Fig. 2C, D). This finding supports theory proposed by
525 Houle (1992); that life history traits may have a higher evolvability than morphological traits.
526 Under the benign environment, CV_A and I_A for fecundity were more than 94% higher than for
527 wing size in both species, and in the stressful environment, fecundity exhibited a CV_A and I_A
528 that was approximately 80% higher than for wing size for *D. serrata* (Fig. 2).

529 The low heritability values detected for fecundity are consistent with classic theory
530 that suggests ultimate fitness traits will exhibit low heritabilities due to directional selection

that fixes beneficial alleles and erodes additive and residual variance (Mousseau and Roff 1987; Falconer and Mackay 1996; Merilä and Sheldon 1999). However, in direct contrast to this, we found that additive variance was actually significantly higher in fecundity where h^2 was low. When examining residual variance ($V_{res} = V_P - V_A$), it becomes evident that increased residual variance is responsible for a reduced heritability in fecundity, rather than eroded additive genetic variance (Table 1; Kruuk et al. 2000; Merilä and Sheldon 2000; McCleery et al. 2004; Moraes et al. 2004). In a study examining how residual and additive variance contributes to heritability values across fitness and morphological traits, Merilä & Sheldon (2000) found fitness traits generally exhibit a higher residual variance compared to morphological traits due to an accumulation of non-additive genetic and early environmental effects. These results support their findings and emphasize the importance of considering trait type when examining how selection shapes additive genetic variance.

An additional aim of this study was to determine whether an easy-to-measure morphological trait can be used as a proxy for fecundity across environments. To examine this, we looked at phenotypic and genetic correlations between fecundity and wing morphology. Although it has been shown that wing length correlates with fecundity in benign environments (Chiang and Hodson 1950; Tantawy and Vetukhiv 1960; Santos et al. 1992; Woods et al. 2002), recent studies have found both evidence for (Woods et al. 2002) or a lack of evidence for (Sgrò and Hoffmann 1998a) positive relationships between wing length, wing width, and fecundity in stressful environments. Here, unadjusted significance tests are suggestive of significant phenotypic correlations between fecundity and wing size in the benign environment for one population of *D. birchii* dams and for *D. birchii* daughters; and in the stressful thermal environment for *D. serrata* daughters. However, these became insignificant after we corrected for False Discovery Rate (Table 2). Although this is a conservative method for multiple comparison in terms of type II errors, the results suggest we

556 cannot use wing size as a proxy for fecundity for these populations in these thermal
557 environments (Fig. 3).

558 Genetic correlations were all fairly low in *D. birchii*, but highly-positive and highly-
559 negative correlations were found in both environments for *D. serrata* (Fig. 3; Supplementary
560 Table 6). Most interestingly in *D. serrata*, fecundity and wing size were significantly
561 negatively-correlated in the benign environment and significantly positively-correlated in the
562 stressful environment. A significant genetic correlation between a pair of traits suggests that
563 the traits are genetically associated through linkage or pleiotropy (influenced by a common
564 locus or loci; Wilson et al. 2010). However, a change in the magnitude or sign of genetic
565 correlations across environments suggests that this genetic association is environment-
566 specific (Falconer and Mackay 1996; see Gutteling et al. 2007 for example). So, while a
567 positive correlation between fecundity and wing size in the stressful environment may
568 indicate that the same gene underlies both traits or the genes influencing both traits are in
569 linkage disequilibrium (Wood and Brodie 2016); a negative correlation in the benign
570 environment may indicate antagonistic pleiotropy between them if this data was looked at
571 independently. However, the drastic change between thermal environments suggests there are
572 environment-specific gene effects that affect these correlations.

573 Additionally, when examining phenotypic correlations and genetic correlations
574 together, we did not find phenotypic correlations that were similar to significant genetic
575 correlations (Fig. 3). This suggests that the environment may be masking phenotypic
576 correlations. The large standard errors associated with many of the genetic correlations also
577 suggest that we may lack sufficient power to detect genetic correlations in some cases. Very
578 large sample sizes are needed in quantitative genetic experiments to estimate heritabilities
579 and genetic correlations with a high degree of precision (Roff 1995; Falconer and Mackay

1996). This is hard to achieve due to logistical challenges, and may partly explain why there is much variation across species, populations, and traits in the literature.

Conclusion

Here, we found that genetic variance and phenotypic and genetic correlations change across thermal environments. However, the direction of these changes was not always consistent across traits, closely-related species, populations within a species, or even generations. This suggests that researchers need to examine adaptive potential specific to their environment, species, and populations if they hope to obtain accurate parameters to predict evolutionary potential. The type of data collected here should represent a starting point for researchers aiming to do so.

Additionally, researchers need to be aware that high genetic variation does not necessarily indicate an increased evolutionary response. Although it is assumed that selection has a strong effect when genetic variation is high and a weak effect when genetic variation is low (when all other factors remain the same), there has been limited evidence showing how selection and genetic variance interact and the studies that have looked at their relationship report a fairly weak association (Wood and Brodie 2016; Ramakers et al. 2018). Future research needs to consider how evolutionary potential is affected by the environment. We show here that genetic variance is highly dependent on temperature and it is accepted that selection is directly mediated by the environment. Yet, specifically in terms of stressful temperatures, a meta-analysis on how selection and genetic variance are coupled found temperature is likely to affect the amount of genetic variation in a population more than the strength of selection (Wood and Brodie 2016). Wood and Brodie (2016) found that temperature affected the amount of genetic variation and the strength of selection in both

605 morphological and fitness traits asymmetrically; meaning the measured impact of
606 temperature stress on genetic variation does not necessarily predict the magnitude of the
607 evolutionary change. Researchers should examine how both genetic variance and selective
608 force (both strength and directionality of selection) is influenced by specific environments to
609 determine the adaptive potential of species to climate change. If a highly positive correlation
610 exists between the two, environmental change would increase both, directly causing
611 increased adaptation; and predictions on how species will adapt to changing environments
612 would be more straightforward (Wood and Brodie 2016; Ramakers et al. 2018; Fischer et al.
613 2020). Genetic correlations also need to be considered in this context. A negative genetic
614 correlation between two traits will constrain evolution on one trait even with an increase in
615 genetic variation and a positive selection differential (and vice versa; Conner 2012; Wood
616 and Brodie 2016). An additional consideration is that the underlying genetic architecture of
617 the trait (polygenic or large-effect loci) should be considered. For example, polygenic traits
618 have been shown to produce greater long-term population viability than in traits affected by
619 large-effect alleles when heritability and the selective force is constant (e.g., Kardos and
620 Luikart 2021). Generally, life-history traits are thought to be polygenic in comparison to
621 large-effect phenotypic traits related to morphology, indicating another reason why trait type
622 needs to be considered when investigating adaptive potential.
623 In conclusion, although we present clear evidence that stressful temperatures affect
624 genetic variation, we did not detect a consistent pattern to that change. These results suggest
625 that adaptive potential cannot be generalized across environments, closely-related species or
626 populations and needs to be considered on a case-by-case basis, specific to the trait in
627 question, and by using a multivariate approach.
628

Acknowledgements

We would like to thank Malaika Chawla, Nicholas Bail, Naomi Laven, Natalie Swinhoe, and Paula Strickland for assistance in data collection, specifically for helping with the task of egg-counting. We thank Jodie Betts for technical assistance. This study was supported by grants from the American Society of Naturalists, the Ecological Society of Australia, the Wet Tropics Management Authority, James Cook University, and the American Australian Association to JC, as well as an Australian Research Council Discovery Grant (DE130100218) to MH.

Conflict of Interest

The authors declare there are no conflicts of interest.

Data archiving

Data will be archived in the Dryad data repository if paper is accepted for publication.

647 **References**

648

649 Adams D, Collyer M, Kaliontzopoulou A (2020) *geomorph: Geometric Morphometric*
650 *Analyses of 2D/3D Landmark Data*. Version 3.3.1.

651

652 Adams DC, Rohlf FJ, Slice DE (2004) Geometric morphometrics: Ten years of progress
653 following the 'revolution'. *Ital J Zool* **71**: 5–16.

654

655 Agrawal AF, Stinchcombe JR (2009) How much do genetic covariances alter the rate of
656 adaptation? *Proc R Soc B Biol Sci* **276**: 1183–1191.

657

658 Altman DG, Bland JM (2011) How to obtain the P value from a confidence interval. *BMJ*
659 **343**: d2304.

660

661 Arnold SJ (1983) Morphology, Performance and Fitness. *Am Zool* **23**: 347–361.

662

663 Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate
664 change on the future of biodiversity. *Ecol Lett* **15**: 365–377.

665

666 Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and
667 Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* **57**: 289–300.

668

669 Blows MW, Hoffmann AA (2005) A reassessment of genetic limits to evolutionary change.
670 *Ecology* **86**: 1371–1384.

671

672 Bubliy OA, Loeschcke V (2002) Effect of low stressful temperature on genetic variation of
673 five quantitative traits in *Drosophila melanogaster*. *Heredity* **89**: 70–5.
674
675 Bubliy OA, Loeschcke V, Imasheva AG (2001) Genetic variation of morphological traits in
676 *Drosophila melanogaster* under poor nutrition: isofemale lines and offspring–parent
677 regression. *Heredity* **86**: 363–369.
678
679 Carreira VP, Soto IM, Mensch J, Fanara JJ (2011) Genetic basis of wing morphogenesis in
680 *Drosophila*: sexual dimorphism and non-allometric effects of shape variation. *BMC Dev Biol*
681 **11**: 32.
682
683 Ceballos G, Ehrlich PR, Dirzo R (2017) Biological annihilation via the ongoing sixth mass
684 extinction signaled by vertebrate population losses and declines. *Proc Natl Acad Sci* **114**:
685 E6089–E6096.
686
687 Charmantier A, Garant D (2005) Environmental quality and evolutionary potential: lessons
688 from wild populations. *Proc R Soc B Biol Sci* **272**: 1415–1425.
689
690 Cheverud JM (1988) A Comparison of Genetic and Phenotypic Correlations. *Evolution* **42**:
691 958–968.
692
693 Chiang HC, Hodson AC (1950) An Analytical Study of Population Growth in *Drosophila*
694 *melanogaster*. *Ecol Monogr* **20**: 173–206.
695

696 Conner JK (2012) Quantitative Genetic Approaches to Evolutionary Constraint: How Useful?
697 *Evolution* **66**: 3313–3320.
698
699 Conner J, Via S (1992) Patterns of phenotypic and genetic correlations among morphological
700 and life-history traits in wild radish, *Raphanus raphanistrum*. *Evolution* **47**: 704–711.
701
702 Czesak ME, Fox CW (2003) Evolutionary Ecology of Egg Size and Number in a Seed
703 Beetle: Genetic Trade-Off Differs Between Environments. *Evolution* **57**: 1121–1132.
704
705 Deutsch CA, Tewksbury JJ, Huey RB, Sheldon KS, Ghalambor CK, Haak DC, *et al.* (2008)
706 Impacts of climate warming on terrestrial ectotherms across latitude. *Proc Natl Acad Sci* **105**:
707 6668–6672.
708
709 Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. John
710 Wiley & Sons, Inc. New York, NY.
711
712 Fischer K, Kreyling J, Beaulieu M, Beil I, Bog M, Bonte D, *et al.* (2020) Species-specific
713 effects of thermal stress on the expression of genetic variation across a diverse group of plant
714 and animal taxa under experimental conditions. *Heredity*: 1–15.
715
716 Fisher RA (1930) *The genetical theory of natural selection*. Clarendon Press: Oxford,
717 England.
718

719 Funk WC, Forester BR, Converse SJ, Darst C, Morey S (2019) Improving conservation
720 policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species
721 Act decisions for conservation practitioners and geneticists. *Conserv Genet* **20**: 115–134.
722
723 Garcia-Gonzalez F, Simmons LW, Tomkins JL, Kotiaho JS, Evans JP (2012) Comparing
724 Evolvabilities: Common Errors Surrounding the Calculation and Use of Coefficients of
725 Additive Genetic Variation. *Evolution* **66**: 2341–2349.
726
727 Gilchrist AS, Partridge L (1999) A Comparison of the Genetic Basis of Wing Size
728 Divergence in Three Parallel Body Size Clines of *Drosophila melanogaster*. *Genetics* **153**:
729 1775–1787.
730
731 Gilchrist AS, Partridge L (2001) The contrasting genetic architecture of wing size and shape
732 in *Drosophila melanogaster*. *Heredity* **86**: 144–152.
733
734 Gómez JM, Abdelaziz M, Muñoz-Pajares J, Perfectti F (2009) Heritability and Genetic
735 Correlation of Corolla Shape and Size in *Erysimum Mediohispanicum*. *Evolution* **63**: 1820–
736 1831.
737
738 Goodall C (1991) Procrustes Methods in the Statistical Analysis of Shape. *J R Stat Soc Ser B*
739 *Methodol* **53**: 285–339.
740
741 Guerra D, Pezzoli MC, Giorgi G, Garoia F, Cavicchi S (1997) Developmental constraints in
742 the *Drosophila* wing. *Heredity* **79**: 564–571.
743

744 Gutteling EW, Doroszuk A, Riksen J a. G, Prokop Z, Reszka J, Kammenga JE (2007)
745 Environmental influence on the genetic correlations between life-history traits in
746 *Caenorhabditis elegans*. *Heredity* **98**: 206–213.
747
748 Hansen TF, Armbruster WS, Carlson ML, Pélabon C (2003) Evolvability and genetic
749 constraint in *Dalechampia* blossoms: Genetic correlations and conditional evolvability. *Mol*
750 *Dev Evol* **296B**: 23–39.
751
752 Hansen TF, Pélabon C, Houle D (2011) Heritability is not Evolvability. *Evol Biol* **38**: 258–
753 277.
754
755 Hausmann NJ, Juenger TE, Sen S, Stowe KA, Dawson TE, Simms EL (2005) Quantitative
756 trait loci affecting delta13C and response to differential water availability in *Arabidopsis*
757 *thaliana*. *Evol Int J Org Evol* **59**: 81–96.
758
759 Higgie M, Blows MW (2008) The Evolution of Reproductive Character Displacement
760 Conflicts with How Sexual Selection Operates Within a Species. *Evolution* **62**: 1192–1203.
761
762 Hoffmann AA, Merilä J (1999) Heritable variation and evolution under favourable and
763 unfavourable conditions. *Trends Ecol Evol* **14**: 96–101.
764
765 Hoffmann AA, Parsons PA (1991) *Evolutionary genetics and environmental stress*. Oxford
766 University Press.
767

768 Hoffmann AA, Schiffer M (1998) Changes in the Heritability of Five Morphological Traits
769 Under Combined Environmental Stresses in *Drosophila melanogaster*. *Evolution* **52**: 1207–
770 1212.
771
772 Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature* **470**:
773 479–485.
774
775 Hoffmann AA, Shirriffs J (2002) Geographic Variation for Wing Shape in *Drosophila*
776 *Serrata*. *Evolution* **56**: 1068–1073.
777
778 Holman L, Jacomb F (2017) The effects of stress and sex on selection, genetic covariance,
779 and the evolutionary response. *J Evol Biol* **30**: 1898–1909.
780
781 Houle D (1992) Comparing Evolvability and Variability of Quantitative Traits. *Genetics* **130**:
782 195–204.
783
784 Kardos M, Luikart G (2021) The genetic architecture of fitness drives population viability
785 during rapid environmental change. *Am Nat*: 37.
786
787 Kellermann VM, van Heerwaarden B, Hoffmann AA, Sgró CM (2006) Very low additive
788 genetic variance and evolutionary potential in multiple populations of two rainforest
789 *Drosophila* species. *Evolution* **60**: 1104–1108.
790

791 Kellermann VM, van Heerwaarden B, Sgró CM, Hoffmann AA (2009) Fundamental
792 Evolutionary Limits in Ecological Traits Drive *Drosophila* Species Distributions. *Science*
793 **325**: 1244–1246.
794
795 Kingsolver JG, Diamond SE, Buckley LB (2013) Heat stress and the fitness consequences of
796 climate change for terrestrial ectotherms. *Funct Ecol* **27**: 1415–1423.
797
798 Kristensen TN, Overgaard J, Lassen J, Hoffmann AA, Sgrò C (2015) Low evolutionary
799 potential for egg-to-adult viability in *Drosophila melanogaster* at high temperatures.
800 *Evolution* **69**: 803–814.
801
802 Kruuk LEB, Clutton-Brock TH, Slate J, Pemberton JM, Brotherstone S, Guinness FE (2000)
803 Heritability of fitness in a wild mammal population. *Proc Natl Acad Sci* **97**: 698–703.
804
805 McCleery R, Pettifor R, Armbruster P, Meyer K, Sheldon B, Perrins C (2004) Components
806 of Variance Underlying Fitness in a Natural Population of the Great Tit *Parus major*. *Am Nat*
807 **164**: E62-72.
808
809 Merilä J, Sheldon BC (1999) Genetic architecture of fitness and nonfitness traits: empirical
810 patterns and development of ideas. *Heredity* **83**: 103–109.
811
812 Merilä J, Sheldon B (2000) Lifetime reproductive success and heritability in nature. *Am Nat*
813 **155**: 301–310.
814

815 Monteiro LR, Diniz-Filho JAF, Reis SF dos, Araújo ED (2002) Geometric estimates of
816 heritability in biological shape. *Evolution* **56**: 563–572.
817
818 Moraes EM, Manfrin MH, Laus AC, Rosada RS, Bomfin SC, Sene FM (2004) Wing shape
819 heritability and morphological divergence of the sibling species *Drosophila mercatorum* and
820 *Drosophila paranaensis*. *Heredity* **92**: 466–473.
821
822 Mousseau T, Roff D (1987) Natural selection and the heritability of fitness components.
823 *Heredity* **59**: 181–97.
824
825 Nguyen TTX, Moehring AJ (2015) Accurate Alternative Measurements for Female Lifetime
826 Reproductive Success in *Drosophila melanogaster*. *PLoS ONE* **10**.
827
828 Orr HA (2009) Fitness and its role in evolutionary genetics. *Nat Rev Genet* **10**: 531–539.
829
830 Overgaard J, Kristensen TN, Mitchell KA, Hoffmann AA (2011) Thermal Tolerance in
831 Widespread and Tropical *Drosophila* Species: Does Phenotypic Plasticity Increase with
832 Latitude? *Am Nat* **178**: S80–S96.
833
834 Pekkala N, Kotiaho JS, Puurtinen M (2011) Laboratory Relationships between Adult
835 Lifetime Reproductive Success and Fitness Surrogates in a *Drosophila littoralis* Population.
836 *PLoS ONE* **6**: e24560.
837
838 *R Core Team* (2019) R: A language and environment for statistical computing. R Foundation
839 for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ramakers JJC, Culina A, Visser ME, Gienapp P (2018) Environmental coupling of heritability and selection is rare and of minor evolutionary significance in wild populations. *Nat Ecol Evol* **2**: 1093–1103.

Reeve ECR (1955) The Variance of the Genetic Correlation Coefficient. *Biometrics* **11**: 357–374.

Robertson FW (1956) Studies in quantitative inheritance XI. Genetic and environmental correlation between body size and egg production in *Drosophila melanogaster*. *J Genet* **83**: 17–32.

Robertson A (1959) The Sampling Variance of the Genetic Correlation Coefficient. *Biometrics* **15**: 469–485.

Roff DA (1995) The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* **74**: 481–490.

Rohlf FJ (2000) On the use of shape spaces to compare morphometric methods. *Hystrix* **11**: 17.

Rohlf FJ (2003) Bias and error in estimates of mean shape in geometric morphometrics. *J Hum Evol* **44**: 665–683.

Rohlf FJ (2010a) *tpsUtil*. Version 1.44. Stony Brook University, NY.

Rohlf FJ (2010b) *tpsRelw*. Version 1.49. Stony Brook University, NY.

Rohlf FJ (2016) *tpsDig*. Version 2.2. Stony Brook University, NY.

Rohlf FJ, Slice D (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst Biol* **39**: 40–59.

Rosenberg A (1982) On the Propensity Definition of Fitness. *Philos Sci* **49**: 268–273.

Rowiński PK, Rogell B (2017) Environmental stress correlates with increases in both genetic and residual variances: A meta-analysis of animal studies. *Evolution* **71**: 1339–1351.

Santos M, Ruiz A, Quezada-Díaz JE, Barbadilla A, Fontdevila A (1992) The evolutionary history of *Drosophila buzzatii*. XX. Positive phenotypic covariance between field adult fitness components and body size. *J Evol Biol* **5**: 403–422.

Saxon AD, O'Brien EK, Bridle JR (2018) Temperature fluctuations during development reduce male fitness and may limit adaptive potential in tropical rainforest *Drosophila*. *J Evol Biol* **31**: 405–415.

Schiffer M, Mcevey SF (2006) *Drosophila bunnanda*-a new species from northern Australia with notes on other Australian members of the montium subgroup (Diptera: Drosophilidae). *Zootaxa* **1333**: 1.

890 Sgrò CM, Hoffmann AA (1998a) Effects of temperature extremes on genetic variances for
891 life history traits in *Drosophila melanogaster* as determined from parent-offspring
892 comparisons. *J Evol Biol* **11**: 1–20.
893
894 Sgrò CM, Hoffmann AA (1998b) Heritable Variation for Fecundity in Field-Collected
895 *Drosophila Melanogaster* and Their Offspring Reared Under Different Environmental
896 Temperatures. *Evolution* **52**: 134–143.
897
898 Sgrò CM, Hoffmann AA (1998c) Effects of stress combinations on the expression of additive
899 genetic variation for fecundity in *Drosophila melanogaster*. *Genet Res* **72**: 13–18.
900
901 Sgrò CM, Hoffmann AA (2004) Genetic correlations, tradeoffs and environmental variation.
902 *Heredity* **93**: 241–248.
903
904 Shaw RG (2019) From the Past to the Future: Considering the Value and Limits of
905 Evolutionary Prediction. *Am Nat* **193**: 1–10.
906
907 Swindell WR, Bouzat JL (2006) Associations between environmental stress, selection
908 history, and quantitative genetic variation in *Drosophila melanogaster*. *Genetica* **127**: 311–
909 320.
910
911 Tantawy AO, Vetukhiv MO (1960) Effects of Size on Fecundity, Longevity and Viability in
912 Populations of *Drosophila pseudoobscura*. *Am Nat* **94**: 395–403.
913

914 Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, *et al.*
915 (2004) Extinction risk from climate change. *Nature* **427**: 145–148.
916
917 Urban MC, Bocedi G, Hendry AP, Mihoub JB, Pe'er G, Singer A, *et al.* (2016) Improving
918 the forecast for biodiversity under climate change. *Science* **353**.
919
920 Walsh B, Blows MW (2009) Abundant Genetic Variation + Strong Selection = Multivariate
921 Genetic Constraints: A Geometric View of Adaptation. *Annu Rev Ecol Evol Syst* **40**: 41–59.
922
923 Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling CA, *et al.* (2010) An
924 ecologist's guide to the animal model. *J Anim Ecol* **79**: 13–26.
925
926 Wood CW, Brodie ED (2015) Environmental effects on the structure of the G-matrix.
927 *Evolution* **69**: 2927–2940.
928
929 Wood CW, Brodie ED (2016) Evolutionary response when selection and genetic variation
930 covary across environments. *Ecol Lett* **19**: 1189–1200.
931
932 Woods RE, Sgrò CM, Hercus MJ, Hoffmann AA (2002) Fluctuating asymmetry, fecundity
933 and development time in *Drosophila*: is there an association under optimal and stress
934 conditions? *J Evol Biol* **15**: 146–157.
935
936 Zelditch M, Swiderski D, Sheets HD, Fink W (2004) *Geometric Morphometrics for*
937 *Biologists: A Primer*. Academic Press: San Diego.
938

939 Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common

940 statistical problems. *Methods Ecol Evol* **1**: 3–14.

941

Figure legends

Figure 1: Parent-offspring quantitative genetic experimental design.

The parent-offspring quantitative genetic experimental design used to measure female fecundity, wing size, and wing shape on both dams and daughters. This design was used for two populations of both *D. birchii* and *D. serrata*. Experimental female flies were raised in either a non-stressful rearing temperature (23°C) or a stressful temperature (28°C). Mass bred populations were raised alongside each generation and supplied males for mating purposes.

Figure 2: Heritability (h^2) and coefficient of additive variance (CV_A) values for a life history trait and morphological traits across a benign (23°C) and stressful (28°C) thermal environment.

Two standardized estimates of additive genetic variance are shown for a life history and two morphological traits in two closely-related species of *Drosophila*. (**A**, **B**) Heritability is standardized by the total genetic variance and (**C**, **D**) coefficient of additive genetic variance is standardized by the trait mean. Evolvability (not shown) will exhibit the same pattern as CV_A. Standard errors (2x) are shown as error-bars, and asterisks indicate significance of the estimate after correction (**.** $P < 0.1$; * $P < 0.05$; *** $P < 0.001$; **** $P < 0.0001$). Standard errors for CV_A were calculated using the standard error estimates from heritability (see Supplementary Tables 1, 2 for details).

Figure 3. **Genetic and phenotypic correlations for fecundity and wing morphology in two sibling-species of *Drosophila*.**

Genetic correlations (r_G) and phenotypic correlations (r_P) between the trait on the x-axis (fecundity and wing size) and the trait on the y-axis (wing size and wing shape RW scores)

967 across a benign and stressful thermal environment in (**A)** *D. birchii* and (**B**) *D. serrata*.
968 Standard errors (2x) for the correlations are indicated by the grey error bars. Asterisks (in
969 red) denote the correlation is significantly different from 0, obtained from the *z*-statistic
970 calculated from standard errors (Altman and Bland, 2011) and *P*-values have been adjusted
971 by the False Discovery Rate method (Benjamini and Hochberg 1995) (* $P < 0.05$; *** $P <$
972 0.001; **** $P < 0.0001$). Phenotypic correlations shown are for the dam generation.

Figure 1

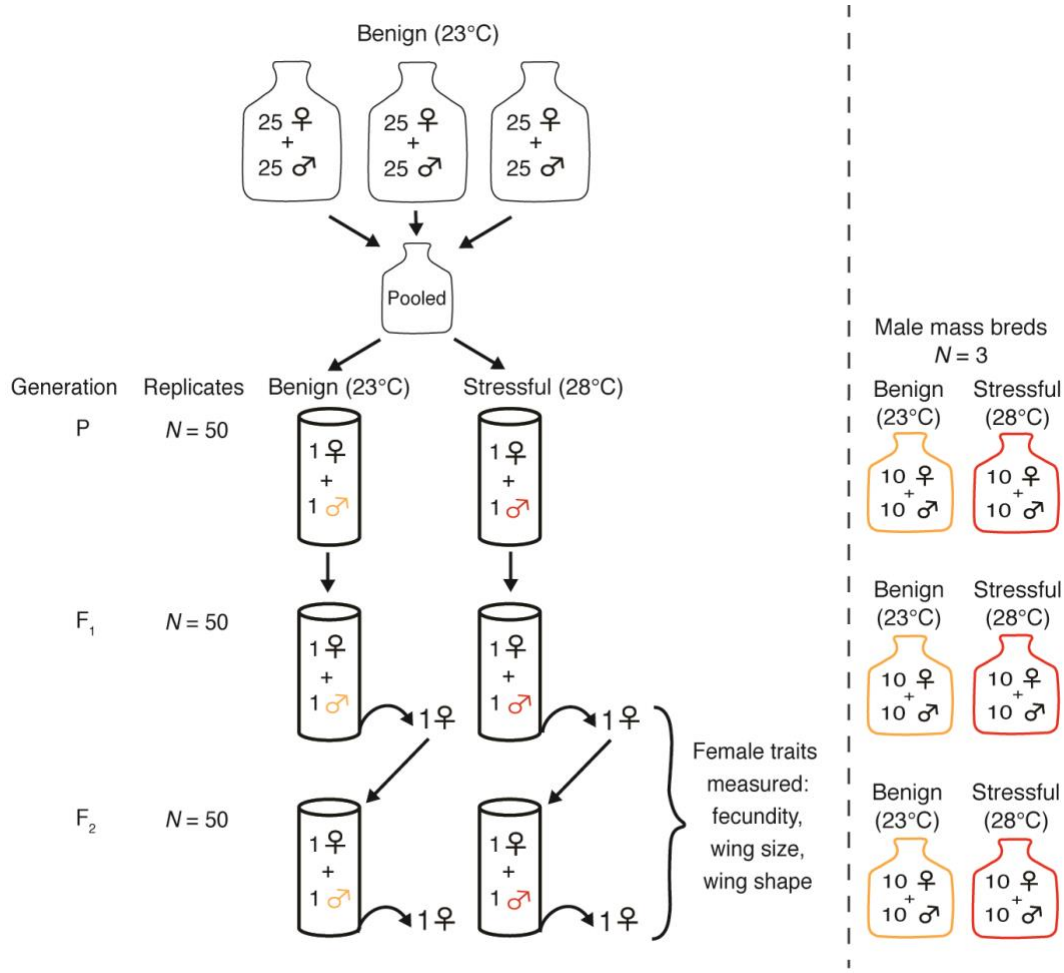

Figure 2

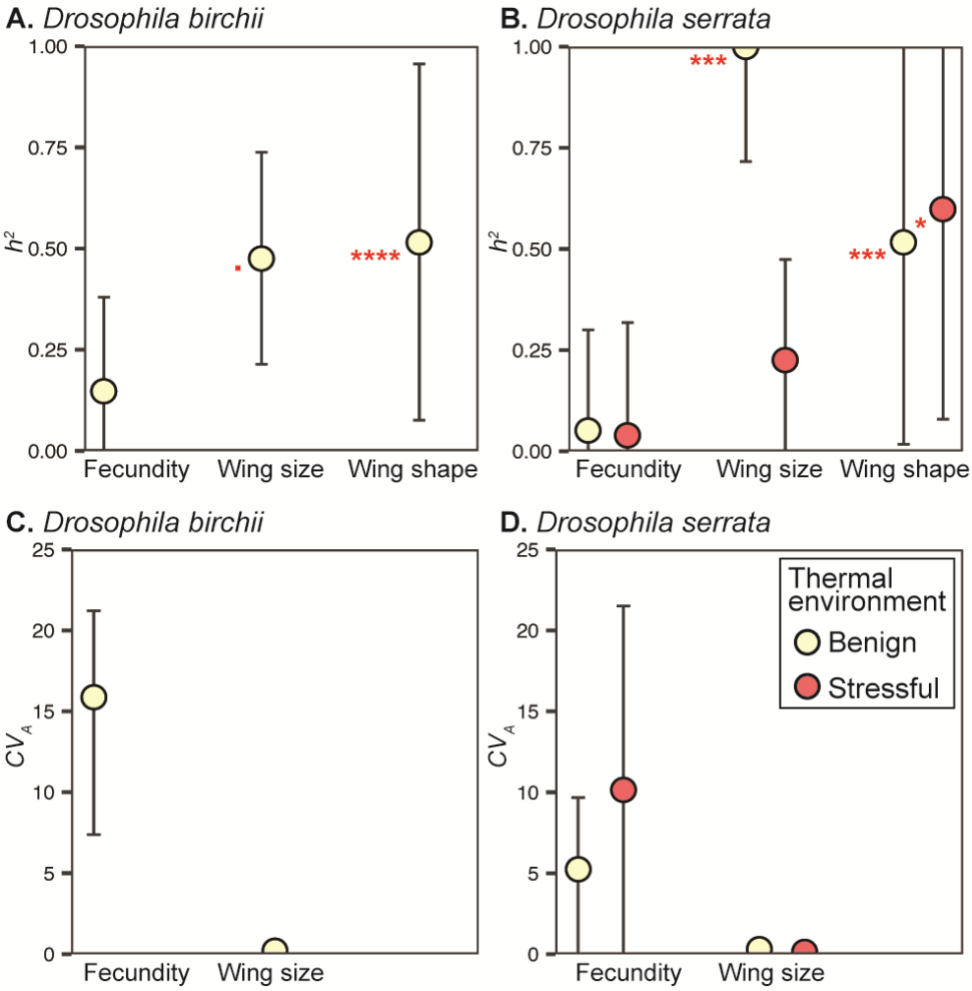

Figure 3

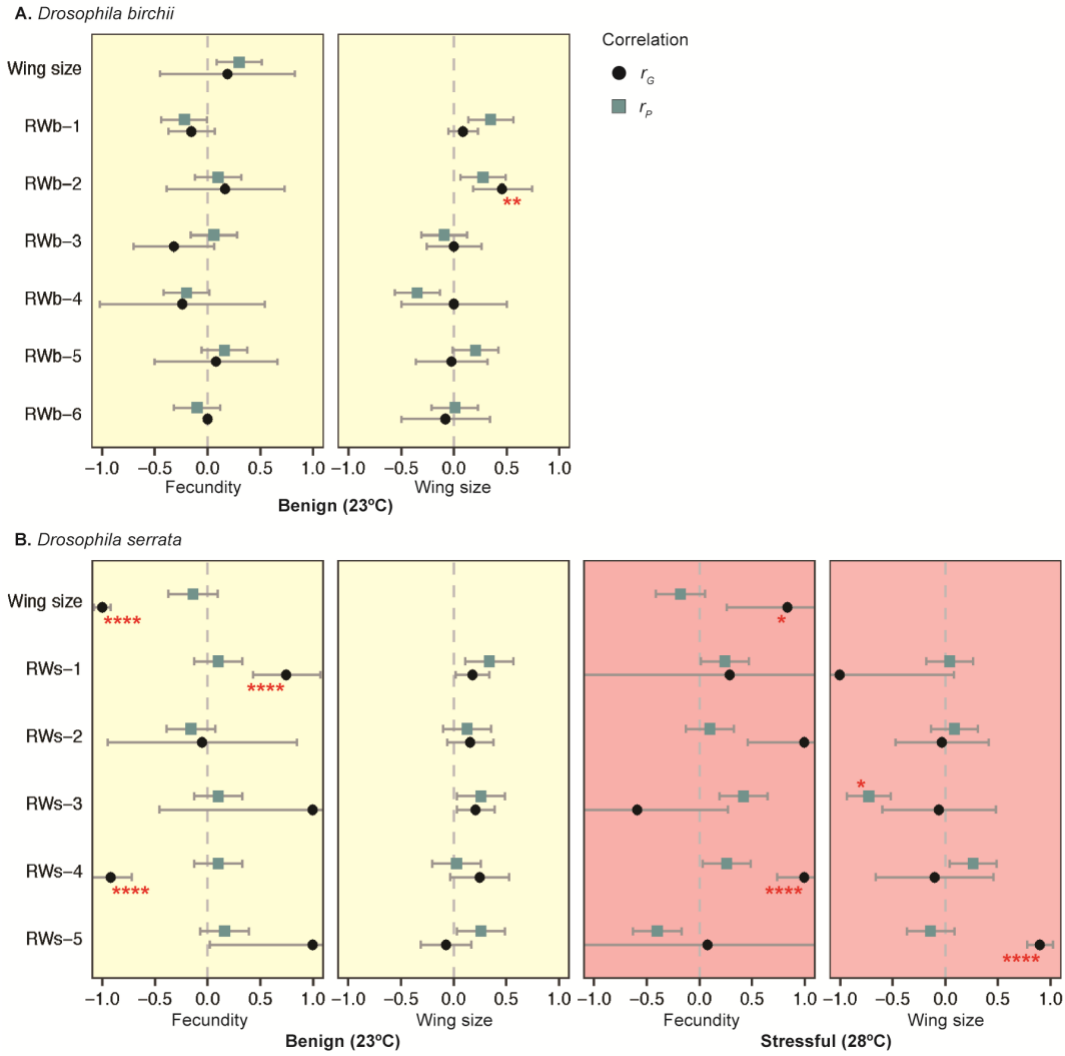

Table 1. Expression of genetic variance parameters for fecundity, wing size, and wing shape; including heritability (h^2), the coefficient of additive variance (CV_A), and evolvability (I_A). Phenotypic (V_P), additive (V_A) and residual (V_{res}) variances are also shown for the pooled dam and daughter values. Population was not a significant contributor to variance, so one metric was calculated per species from parent-offspring regressions. Bold values indicate a slope significantly different than zero and asterisks indicate significance level after correction for False Discovery Rate (**.** $P < 0.1$; * $P < 0.05$; *** $P < 0.001$; **** $P < 0.0001$). Parameters could not be calculated for *D. birchii* within the stressful environment because daughters did not develop. CV_A and I_A values shown are x 10^2. Values for individual relative warp scores for wing shape can be found in Supplementary Table 3.

		Benign (23ºC)							**Stressful (28 ºC)**						
Trait	Species	*N*	$h^2 \pm SE$	V_P	V_A	V_{res}	CV_A	I_A	*N*	$h^2 \pm SE$	V_P	V_A	V_{res}	CV_A	I_A
Fecundity	*D. birchii*	86	0.148 ± 0.116	1384.6	204.92	1179.68	15.88	2.524	-	-	-	-	-	-	-
	D. serrata	69	0.052 ± 0.124	1105.1	57.47	1047.64	5.26	0.276	65	0.040 ± 0.139	1879.5	75.18	1804.32	10.17	1.032
Wing size	*D. birchii*	81	**0.476** ± 0.131**.**	0.0005	0.0002	0.0003	0.224	0.0005	-	-	-	-	-	-	-
	D. serrata	64	**1.000** ± 0.142***	0.0005	0.0005	0.0000	0.324	0.0011	62	0.226 ± 0.124	0.0005	0.0001	0.0004	0.156	0.0002
Wing shape	*D. birchii*	81	**0.516** ± 0.22****	-	-	-	-	-	-	-	-	-	-	-	-
	D. serrata	64	**0.517** ± 0.25***	-	-	-	-	-	62	**0.599** ± 0.26*	-	-	-	-	-

Table 2: **Phenotypic correlations between fecundity and wing size.** r_P is the phenotypic correlation and the *P*-values were obtained from an *F*-test of the linear regression of one trait on the other and both unadjusted (raw) and adjusted (corrected for using the False Discovery Rate method (Benjamini and Hochberg 1995) are shown. Sample sizes (*N*) indicate the number of individuals used in each correlation. Phenotypic correlations were calculated for each population separately when population was found to be a significant contributor to variation.

		Benign (23ºC)				**Stressful (28 ºC)**			
Species	Generation Population	*N*	r_P	*P*-value (raw)	*P*-value	*N*	r_P	*P*-value (raw)	*P*-value
D. birchii	Dams	86	0.30	0.168	0.437	78	0.22	0.346	0.647
	Mt. Lewis	45	-0.14	0.655	0.763	-	-	-	-
	Paluma	41	0.76	0.014*	0.104	-	-	-	-
	Daughters	87	0.58	0.007**	0.073	-	-	-	-
D. serrata	Dams	78	-0.14	0.526	0.760	77	-0.18	0.422	0.707
	Daughters	67	0.12	0.628	0.763	65	0.52	0.035*	0.202