Genetic variance in fitness and its cross-sex covariance predict adaptation during experimental evolution.

Short running title: Adaptive potential to new environments

Authors: Eva L. Koch[1,2], Sonja H. Sbilordo[1], and Frédéric Guillaume[1]

Affiliations: [1] Department of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland
[2] Department of Animal and Plant Science, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom

Corresponding author:
Frédéric Guillaume
phone: +41 44 635 66 23
guillaume.fred@gmail.com

Abstract

In presence of rapid environmental changes, it is of particular importance to assess the adaptive potential of populations, which is mostly determined by the additive genetic variation (V_A) in fitness. In this study we used *Tribolium castaneum* (red flour beetles) to investigate its adaptive potential in three new environmental conditions (Dry, Hot, Hot-Dry). We tested for potential constraints that might limit adaptation, including negative genetic covariance between female and male fitness. Based on V_A estimates for fitness, we expected the highest relative fitness increase in the most stressful condition Hot-Dry and similar increases in single stress conditions Dry and Hot. High adaptive potential in females in Hot was reduced by a negative covariance with male fitness. We tested adaptation to the three conditions after 20 generations of experimental evolution and found that observed adaptation mainly matched our predictions. Given that body size is commonly used as a proxy for fitness, we also tested how this trait and its genetic variance (including non-additive genetic variance) were impacted by environmental stress. In both traits, variances were sex and condition dependent, but they differed in their variance composition, cross-sex and cross-environment genetic covariances, as well as in the environmental impact on V_A.

Keywords: fundamental theorem of natural selection, non-additive genetic effects, heritability, quantitative genetics, evolvability

Introduction

Environmental changes pose a substantial risk of extinction to many organisms (Thomas et al. 2004; Parmesan 2006). Predicting whether a population is able to persist is therefore of crucial importance. Species may adapt via plastic or genetic changes. Plastic changes, e.g. physiological or behavioural adjustments, allow individuals to cope with stressful conditions (Charmantier et al. 2008). However, plastic responses are often costly and thus likely limited (Houle 1992; DeWitt et al. 1998; de Jong 2005; Valladares et al. 2007; Pfennig et al. 2010; Snell-Rood et al. 2010; Sokolova et al. 2012). It is also not clear whether they are sufficient to compensate negative effects of environmental changes (Duputié et al. 2015; Arnold et al. 2019). Instead, fast environmental shifts may require rapid adaptation by genetic evolution on short time scales. In this case, the standing genetic variation already present in the population is of particular importance (Kellermann et al. 2006; Bell and Gonzalez 2009). Adaptive evolution proceeds through different individual contributions to the next generation, i.e. differences in their fitness. Importantly, this variation in individual fitness has to be at least partly due to some underlying genetic variation in order to change the genetic composition of a population. Usually, only variance due to additive genetic effects (V_A) is considered since additive effects are inherited from parents to offspring and determine the response to selection. V_A of relative fitness gives the expected increase of fitness in the next generation (Fisher 1930; Price 1972; Falconer and MacKay 1996). Thus, the existence of V_A in fitness is a prerequisite for adaptation and can be used as an overall estimate of a population's capacity to adapt, or its evolvability (Houle 1992; Hansen et al. 2011; Shaw and Shaw 2014).

A common and intuitive expectation is that V_A in fitness should be low because selection depletes genetic variation in adapted populations. Accordingly, it was found that heritability (h^2, proportion of V_A in total

variance V_P) was lower in traits more closely associated with fitness (Mousseau and Roff 1987; Kruuk et al. 2000; Merilä and Sheldon 2000; Teplitsky et al. 2009; Wheelwright et al. 2014). However, h^2 is not ideal for estimating V_A in fitness and the evolutionary potential of a population (Hansen et al. 2011; Wheelwright et al. 2014; Morrissey and Bonnet 2019) since low h^2 is often due to higher environmental variance (Schluter et al. 1991; Merilä and Sheldon 1999). Many populations may also not be at an evolutionary equilibrium (Shaw and Shaw 2014) in environments that have been recently changed by human impacts (Fugère and Hendry 2018). Estimates of V_A in fitness are scarce, especially in natural populations and tend to show large heterogeneity in the estimates of V_A for lifetime fitness (e.g., in vertebrate populations: Kruuk *et al.*, 2000; Merilä & Sheldon, 2000; McCleery *et al.*, 2004; McFarlane *et al.*, 2014). Although it is predicted that some species might not have the evolutionary potential to adapt to shifting environmental conditions (Etterson and Shaw 2001), some studies found substantial amounts of V_A in fitness in wild populations. Similarly, laboratory studies have reported significant V_A for fitness related traits in *D. melanogaster* (Gardner et al. 2005; Fry 2008; Long et al. 2009).

Despite existing V_A in fitness traits, evolution to new conditions can be constrained by antagonistic pleiotropy if alleles influence several fitness components but with opposite effects. This leads to trade-offs since one component cannot be optimized without reducing the other, e.g. fecundity and life length (Roff 2000). Another example of trade-offs is sexual antagonism (Foerster et al. 2007; Bonduriansky and Chenoweth 2009; Delcourt et al. 2009; Kirkpatrick 2009; Poissant et al. 2010; Calsbeek et al. 2015; Connallon and Hall 2016). Fitness optima might often differ between males and females. However, sharing a great part of their genome constrains independent evolution and limits adaptation when selection in the two sexes is opposite to the genetic correlation. Such a constraint should be revealed by negative genetic correlations between male and female fitness. Similar constraints may appear when adaptation to certain environmental conditions (e.g., elevated temperature) trades-off with adaptation to other conditions. Environmental changes often include

simultaneous changes in several variables. For short-term stress exposure, effects like cross-tolerance and hardening (i.e., resistance to one stress develops after exposure to another stress) were observed, and it is known that some stress responses rely on the same physiological mechanisms (Bubliy et al. 2012). Evolutionary adaptation to different stressors involving the same pathways may thus lead to correlated resistance to another stressor (Bubliy and Loeschcke 2005; Sikkink et al. 2015). However, examples of local adaptation suggest that selection may often favour different genotypes in different conditions (Hereford 2009). Immediately after exposure to an environmental change, genetic correlations between fitness in different conditions can inform us to what extent the genetic basis of fitness is shared between environments. If we observe negative genetic covariances between fitness in different conditions, it is likely that alleles providing fitness benefits in one condition become detrimental in another. Adaptation when both stress factors are experienced at the same time can then be limited.

Changes in environmental conditions are also well known to affect genetic variances (Sgrò and Hoffmann 1998; Hoffmann and Merilä 1999; Rowiński and Rogell 2017), and covariances (Simons and Roff 1996; Sgrò and Hoffmann 2004; Wood and Brodie 2015), including cross-sex genetic covariance (Delcourt et al. 2009; Poissant et al. 2010; Punzalan et al. 2014). This can have substantial implications for evolutionary potential (Wilson et al. 2006; Husby et al. 2011), since the environmental shift that imposes a risk of extinction can at the same time either increase V_A for fitness (Shaw and Shaw, 2014) or reduce the evolutionary potential to adapt to this new condition (Wood and Brodie 2016). To fully understand the impact of environmental change on a population's persistence it is therefore essential to know how genetic variances change in different environments and to identify potential constraints.

The fundamental importance of V_A of fitness for predicting contemporary evolution (Shaw and Etterson 2012; Hendry et al. 2018) and recent statistical advances in quantitative genetics have fostered great interest in estimating the adaptive capacity of wild populations (Charmantier et al. 2014). Even more so with the recent

progress in using genomic markers to infer genetic resemblance among individuals (Gienapp et al. 2017; Perrier et al. 2018). These advances open up new perspectives for applications of classical quantitative genetic and genomic tools in wild populations, addressing important questions regarding populations' persistence under environmental change (Waldvogel et al. 2020). However, so far relatively little information exists about the predictive value of V_A for fitness over several generations. In our study, we addressed this question by combining classical quantitative genetics with experimental evolution in the model organism *Tribolium castaneum*. We used a two-generation half-sib/full-sib breeding design to estimate genetic variances of fitness traits in four different conditions (control, dry, hot and hot-dry) representing two often co-occurring stressors, heat and drought. We measured offspring number as estimate of fitness in the F1 generation, and body size in the F2 generation, as an additional trait, often use as fitness proxy. We evaluated adaptation to heat, drought and their combination after 20 generations of experimental evolution. Thus, our experimental setup allowed us to explore many different facets of adaptation and to ask: how the adaptive potential changes under stressful conditions, whether trade-offs between female and male fitness can constrain adaptation, and by testing for genotype-by-environment interactions (G x E,) to which extent resistance to different stressors shares a common genetic basis. Having obtained such estimates in the founder populations, we could then gain deeper insights into the process of adaptation. Linking experimental evolution and classical quantitative genetics proved to be a powerful approach to evaluate the predictive power of genetic variance for fitness, and obtain a better understanding of the observed adaptation after 20 generations in new environments.

Material and Methods

Animal rearing and stress treatments

We used the *Tribolium castaneum* Cro1 strain (Milutinović et al. 2013), an outbred lab strain collected from a wild population in 2010, kept at high population size (>10,000) and adapted to lab standard conditions (33°C, 70% relative humidity) for more than 30 generations. Beetles were kept in 24h darkness on organic wheat flour mixed with 10% organic baker's yeast. We sterilized the flour and yeast by heating them for 12h at 80°C before use. We measured fitness as well as size in control (=standard) conditions and three stress treatments

with increased temperature or/and decreased humidity. The conditions in the treatments were: Hot: 37°C and 70% relative humidity, Dry: 33°C and 30% r. h., Hot-Dry: 37°C and 30% r. h.

In order to be able to estimate genetic variances, we applied a split-brood paternal half-sib breeding design. We produced 147 half-sib families by mating virgin males to three virgin females (Figure S1A). Half- as well as full-sib families were split across all conditions (Figure S1A). Male and female offspring (four females and two males per full-sib family and condition) were separated at the pupal stage and transferred to 10 mL tubes with 1 g of medium and remained there until they were used for the fitness assay eight weeks later.

Fitness Assay and producing double first cousins

To estimate fitness, we mated each virgin male with two unrelated virgin females from the same condition in 15mL tube with 1g medium. The male was removed after 24h and females transferred into two separate tubes. Females were removed from the tubes after one week of egg laying, and 9g medium was added to provide food for the developing offspring. After five weeks the number of adult offspring was counted. While we conducted the matings for the fitness assay, we followed a specific crossing design and always crossed two pairs of full-sib families (Figure S1B). Individuals resulting from these crosses (F2) were double first cousins.

Size

Body size was measured in the F2, i.e. in the offspring of beetles that were used for the fitness assay. To estimate body size, we used the centroid size of the abdominal segment IV as proxy for total size since it can be measured more accurately than dry weight in very small insects and shows a high correlation with body mass (Wickman and Karlsson 1989; Honěk 1993). We decided to use size of the abdominal segment because we could measure it with higher accuracy than the total size, which strongly depends on whether an individual is in perfectly stretched position. The size of other parts of the body show a high correlation with this segment (Supplemental Figure S3). Dead beetles were fixed with a double-faced Scotch tape dorsally on a microscope slide. Sex was determined based on the sex patches of males at the inside of the femur. After sexing, all legs of

the beetles were removed. Slides with single specimens were placed under a Wild M8 Heerbrugg M8 dissection microscope with a transmitted light stand. Two further light sources from above were installed for better illumination. Images were captured with a 25x magnification using a Leica DFC495 digital camera connected to a PC running the Microscope imaging software LAS v4.6.2. File utility program "tpsUtil" was used to build tps files from the images and tpsdig2 for setting the landmarks (see http://life.bio.sunysb.edu/ee/rohlf/software.html for program information). To estimate beetle size, four landmarks were set on the ventral part of the abdominal IV (https://figshare.com/articles/_Morphology_of_the_red_flour_beetle_Tribolium_castaneum_/759706/1). The centroid size was calculated using the free software Past 3.14 (for information see: http://folk.uio.no/ohammer/past/).

Effects of condition on fitness and size

Statistical analysis was conducted in R version 3.4.2 (R Core Team 2017). We used linear mixed models as implemented in the Rpackage *lme4* 1.1-17 (Bates et al. 2015) to explore the effects of treatments on fitness and size. For fitness analysis we included treatment and batch (time of fitness assay) as fixed effects and mother identity as random effect to account for non-independence since some of our measured individuals were full-sibs. For analysing size we included batch (individuals that grow up at the same time), treatment and sex as well as their interaction as fixed effects and mother identity as random effect. Significance and confidence intervals were obtained using the Rpackages *lsmeans* 2.27-62 (Lenth 2016) and *lmertest* 3.0-1 (Kuznetsova et al. 2017).

Adaptation

We used ten replicate lines per condition originating from the same ancestral population (same as before, Cro1) and let them adapt for 20 generations (Supplemental Figure S2). Each new generation was set up by randomly selecting 120 pupae and placing them into a new vial with 70g medium. One selection line in Dry

176 became extinct. Adult beetles of generation 20 from all selection lines were transferred to control conditions,
177 in which they stayed for one week to mate and lay eggs. After removal of the adults, we waited until their
178 offspring had reached the pupal stage and separated males and females. These individuals (generation 21)
179 developed completely in control conditions. When they had reached the adult stage, each virgin male was
180 mated with a virgin female of the same selection line and their offspring was transferred to all four conditions
181 in the egg stage, resulting in full-sib families split across all conditions (Figure S2). As soon as these offspring
182 (generation 22) had reached the pupal stage, males and females were separated. To compare fitness of
183 different selection lines and test for adaptation, a virgin male and a virgin female of the same selection line in
184 the same condition, but from different families were mated and the number of adult offspring produced within
185 four days of mating and egg laying was used as a fitness estimate. To test for adaptation, we compared
186 whether offspring number of selection lines in their native condition was significantly higher compared to non-
187 adapted control lines. First, we analysed each condition separately. We used linear mixed models including
188 selection regime as fixed effect and lines and families nested within lines as random effects using the
189 Rpackage *lme4* (Bates et al. 2015). We were further interested to test for correlated responses, i.e. whether
190 adaptation to a certain stress treatment could increase fitness in another. For this we run one analysis using
191 the complete data set. Selection regime, conditions, and their interaction were used as fixed effects, lines,
192 families nested within lines, and line-treatment interaction as random effects. Reported effect sizes and
193 standard errors were obtained from the summary output of the model. p-values and confidence intervals were
194 computed with the Rpackages *lmertest* (Kuznetsova et al. 2017) and *lsmeans* (Lenth 2016). Statistical analyses
195 were conducted in R (R Core Team 2017). Model diagnostic plots are shown in supplemental Figure S6 and the
196 complete results in Table S2.
197
198 **Quantitative genetic analyses**
199 We estimated genetic variances, covariances and correlations using an animal model and restricted maximum
200 likelihood estimation as implemented in Asreml version 3.0 (Gilmour et al. 2009). The animal model is a linear
201 mixed effect model that uses all known relationships from a pedigree as random effect to partition observed

variance into additive genetic variance and other sources of variance (Kruuk 2004). Non-additive genetic relationship matrices were created with the Rpackage *nadiv* 2.16.0.0 (Wolak 2012). All models were run in R by Asreml-R (Butler et al. 2009). All models reached convergence.

Fitness

First, we ran a series of univariate models for each condition separately with offspring number as response variable, batch (samples where fitness assay was started on the same day) as fixed effect, additive genetic effects of females and additive genetic effects of males as random effects. Maternal identity (mother of female) was included as random effect to account for resemblance of full-sibs due to maternal, common environment or non-additive genetic effects. Significance of random effects was determined by likelihood ratio tests, testing whether excluding a certain random effect resulted in a significantly worse model. Reported P-values are one-tailed since we tested whether a certain variance is different from zero and variances cannot become negative (Wilson et al. 2010). Although count data, it can be analysed assuming a Gaussian distribution when distribution converges towards a normal distribution (de Villemereuil 2018). Distribution of offspring numbers are shown in Figure S4. Model diagnostic plots and more details are given in the Supplement (Figure S5). For estimating cross-sex genetic correlation we used the same univariate models but included a covariance between female and male additive genetic effects. Maternal effects were removed in subsequent analyses since they were very small and non-significant. Studies so far used bivariate models to investigate cross-sex genetic correlations (Brommer et al. 2007; Foerster et al. 2007; McFarlane et al. 2014; Punzalan et al. 2014; Wolak et al. 2018). In case of fitness this could be problematic, because the same observation (number of adult offspring resulting from a mating) would be used twice. Using univariate models including additive genetic effects of females and males avoids pseudoreplication. Significance of genetic correlations was determined by two-tailed likelihood ratio tests comparing a model including correlations to a model with correlations set to zero. In order to test for genotype-by-sex interaction (G x S), which indicates that a genotype is differently expressed in males and females, we compared the model to a constrained model

with genetic correlation fixed to one. To examine whether additive genetic effects of females and males were different, we tested whether the unconstrained model was significantly better than a model with V_A of females and V_A of males forced to be equal within each condition.

For estimating cross-condition genetic covariances and correlations we used bivariate models with fitness in two conditions as response, batch as fixed and additive genetic effects of females and males as random effects. Since each individual was only measured in one condition, residual covariance was set to zero. Significance of correlation was assessed by comparing to a model with correlation fixed to zero using two-tailed likelihood ratio tests. Genetic correlations between traits measured in different environments that are significantly less than unity, indicate a G x E interactions (Kruuk 2004; Charmantier and Garant 2005). Comparison with a model with correlation fixed to one was used to test for G x E (Wilson et al. 2010). In all bivariate models, covariance between male effects was set to zero, because male effects were very small and not significant in some conditions and hence it was not meaningful to estimate covariances.

To allow comparisons across different conditions, we calculated h^2 and CV_A (CV_A is the square root of V_A divided by the phenotypic mean of the trait, multiplied by 100) of fitness for males and females in each condition. I_A (V_A divided by trait mean squared, multiplied by 100 (Houle 1992)) was calculated as an estimate for the proportional change after one generation. Standard errors of CV_A and I_A were computed as described in Garcia-Gonzalez *et al.*, (2012).

To assess whether genetic variances were significantly affected by environmental change, we used a multivariate model with fitness in each condition as separate response variable. Since fitness between conditions differed in mean and variance, we standardized data. We applied two different standardizations: 1) mean standardized (i.e. relative fitness) and 2) dividing by the standard deviation (sd) to have a variance of one. We then tested whether constraining the multivariate model and forcing V_A of standardized fitness to be equal in all conditions resulted in a significantly worse model. If the constrained models are worse, we can conclude that mean standardized V_A (I_A of absolute numbers), or sd standardized V_A (h^2 of absolute numbers) are significantly influenced by the environment. Similarly, we tested whether genetic cross-condition correlations were different. We used a multivariate model and constrained all pairwise cross-condition

253 correlations to be equal. We then tested if this model was significantly worse than an unconstrained model
254 using a two-tailed likelihood-ratio test.
255
256 *Size*
257 Our crossing design during the fitness assay (reciprocal crossing of two pairs of full-sibs, see Figure S2) allowed
258 us to estimate non-additive genetic variance (V_D) in the following generation (F2). In a paternal half-sib
259 breeding design these effects contribute to resemblance among full-sibs and cannot be separated from
260 maternal or common environmental effects. In contrast, double first cousins (DFC) share non-additive genetic
261 effects without being confounded by maternal effects or a common environment. We partitioned the
262 observed variance for size into additive (V_A), maternal (V_M), non-additive (V_D), and residual variance V_R while
263 controlling for batch effects. Batches represent individuals that grew up at the same time, and thus accounts
264 for variations in the medium or lab temperature. Similar to the analysis of fitness data, we first analysed each
265 condition separately using univariate animal models with batch as fixed and maternal, additive, and non-
266 additive genetic effects as random effects. To get sex specific estimates, we analysed male and female size
267 separately. Genetic correlations between sexes and between conditions were assessed by bivariate models.
268 Significance of correlations, G x E, environmental effects on I_A and h^2 of size were tested in the same way as we
269 described before.
270
271 To demonstrate how responses are affected by covariances between the sexes we used estimated variances
272 and covariances between male and female additive genetic effects of relative fitness to predict fitness increase
273 after one generation. We applied the multivariate breeder's equation $\Delta z = G\beta$ (Lande 1979), where Δz is a
274 vector of changes in the trait means, G the 2 x 2 genetic variance-covariance matrix for female and male
275 relative fitness estimated in each condition, and β a vector of selection gradients. We used a vector of
276 selection gradients that assumed equal selection on male and female fitness, β^T=[1,1]. Total change in fitness is
277 then the sum of fitness increase in females and in males. Alternatively, we may consider male additive genetic
278 effects as indirect genetic effects on female reproductive output. Joint effects of direct and indirect selection

on evolutionary changes of a single trait can be calculated as described in Bijma & Wade (2008) (equation 14): $\sigma_T^2 = \sigma_{A_D}^2 + 2(n-1)\sigma_{A_{DI}}^2 + (n-1)^2\sigma_{A_I}^2$, where σ_T^2 is the total heritable variance, $\sigma_{A_D}^2$ additive genetic variance for direct effects (female V_A), $\sigma_{A_I}^2$ additive genetic variance for indirect effects (male V_A), and n gives the number of interacting individuals (here, it equals two). In case of relative fitness, this total heritable variance is equal to the predicted relative increase in fitness (Fisher 1930). Thus, considering male and female fitness as two separate traits and applying the multivariate breeder's equation, or treating additive genetic effects of males as indirect genetic effects on female fitness is equivalent.

Results

Effect of treatment

It was previously shown that all treatments had a highly significant effect on offspring number (Koch and Guillaume 2020) with the effect of heat being stronger and the lowest offspring number when heat and drought were combined. (Figure 1A). Treatments also showed a significant effect on size ($F_{3,7710}$= 76.30, p < 2.20E-16, Table S1) and size decreased in the stressful conditions. Similar to offspring number, the effect of drought was smaller. Combining both stressors did not result in an additional decrease in size. In contrast, we observed the lowest size in Hot and higher sizes in Hot-Dry (Figure 2A). The difference between males and females was significant ($F_{1,7759}$ = 518.19, p < 2.20E-16) with females being larger. We also detected a significant sex-by-condition interaction ($F_{3,7759}$ = 7.27, p = 8.16E-04) indicating that males were more sensitive to stressful conditions (Figure 2A).

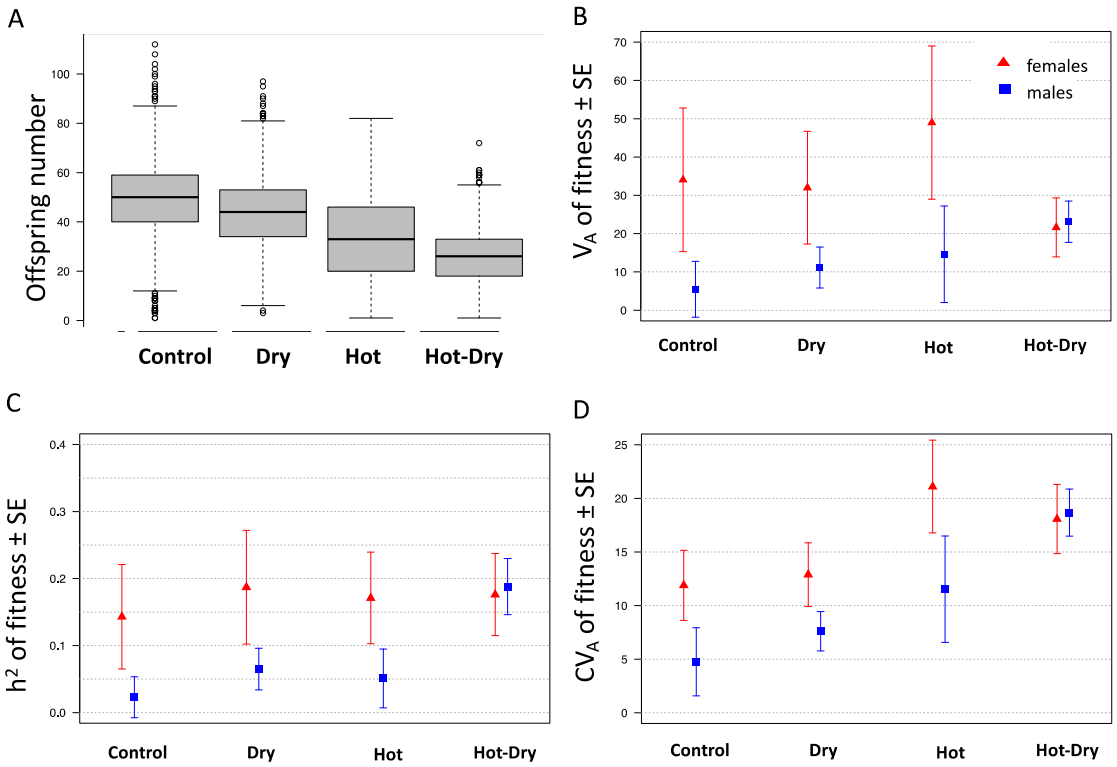


Figure 1: Fitness (= number of adult offspring) **(A)** and estimates for its additive genetic variance (V_A) **(B),** heritability (h^2) **(C),** and coefficient of V_A (CV_A), which is the square root of V_A divided by the phenotypic trait mean, multiplied by 100 **(D)** in female and male flour beetles (*Tribolium castaneum*) in four different environmental conditions.

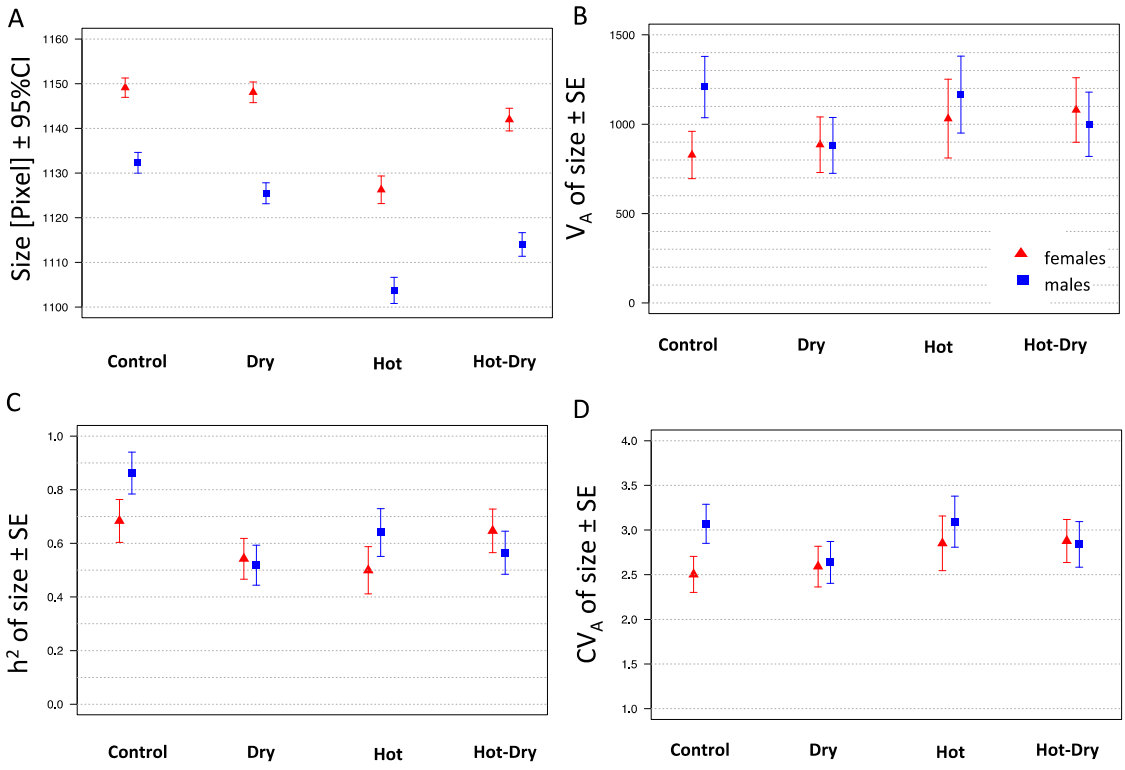


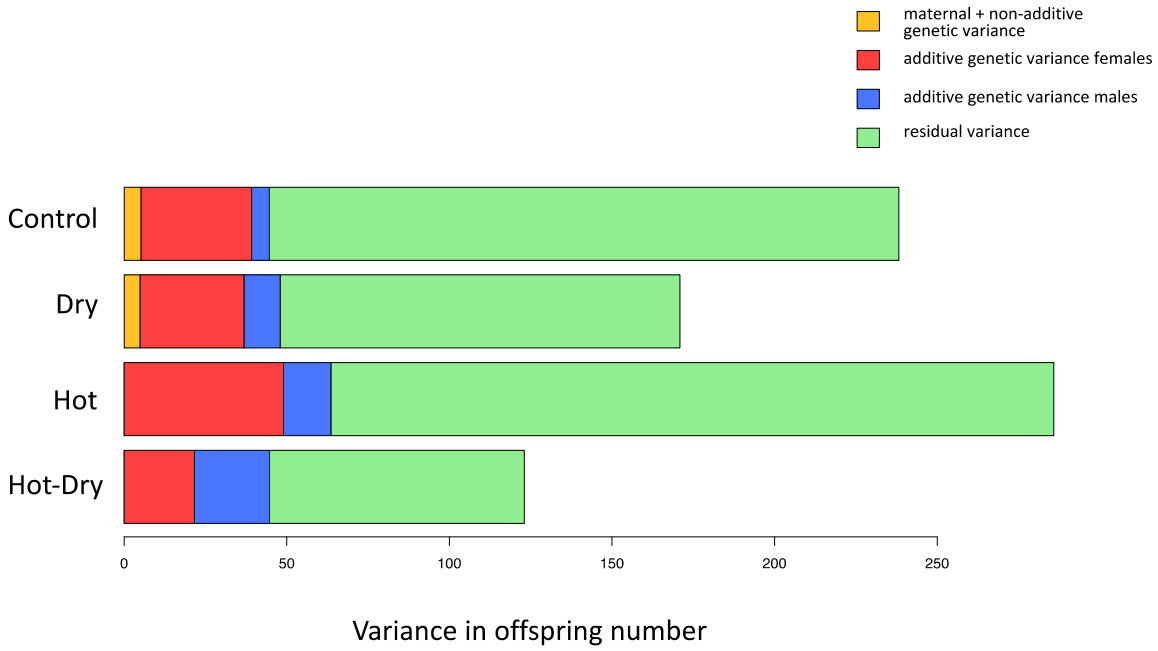
Figure 2: Body size (centroid size of abdominal segment IV) **(A)** and estimates for its additive genetic variance (V_A) **(B),** heritability (h^2) **(C)**, and coefficient of V_A (CV_A) **(D)** in female and male flour beetles (*Tribolium castaneum*) in four different environmental conditions.

Genetic variances and covariances

Fitness

V_M, which consisted of variance due to maternal effects, common environment and non-additive effects, was not significant in any condition. We found significant V_A for female fitness in all conditions (Table 1). V_A for males was generally smaller (P = 8.10E-03) and not significantly different from zero in Control and Hot. A remarkable exception was Hot-Dry, where male V_A was increased and of the same magnitude as female V_A (Figure 3). V_A in females was similar in Control and Dry but increased in Hot and was reduced in Hot-Dry (Figure 1B). Since residual variances changed simultaneously with V_A (Table 1), h^2 for females was similar in all stress conditions (Figure 1C). When we compared genetic variances using standardized data, we found h^2 of males (P

316 = 2.75E-03), but not females (P = 0.55) to differ significantly between conditions. Similarly, I_A was affected by
317 conditions in males (P = 8.87E-05), but not in females (P = 0.07). When we tested a model constraining V_A of
318 females and males to be the same in all conditions, we found it to be significantly worse (P = 2.46E-08) than an
319 unconstrained model, thus showing that the total amount of heritable genetic variance (V_A of females and
320 males) was different between treatments.



321 **Figure 3:** Variance components of offspring number of *Tribolium castaneum* in four different environmental conditions
322 estimated using univariate animal models.

Table 1: Genetic variances (V_A: additive genetic variance; V_M: maternal variance; V_R: residual variance), heritability (h^2), coefficient of additive genetic variance (CV_A) of offspring number and cross-sex additive genetic covariances (COV_A) and correlations (r_A) in different environmental conditions. Estimates for genetic variances were obtained from a univariate animal model including additive genetic effects of males and females as random effects. For COV_A and r_A univariate animal models were used with covariance between female and additive genetic effects. N gives the number of reproducing females used for the analysis. All significant results are in bold.

	Condition			
	Control	**Dry**	**Hot**	**Hot-Dry**
N	1514	1603	1005	1396
mean (SE)	49.09 (0.4)	43.88 (0.36)	33.18 (0.55)	25.73 (0.31)
V_A females (SE)	**34.06 (18.75)**	**32.00 (14.71)**	**48.99 (20.02)**	**21.62 (7.71)**
h^2 females (SE)	**0.14 (0.08)**	**0.19 (0.09)**	**0.17 (0.07)**	**0.18 (0.06)**
CV_A females (SE)	**11.89 (3.27)**	**12.89 (2.97)**	**21.10 (4.32)**	**18.07 (3.23)**
V_A males (SE)	5.44 (7.29)	**11.14 (5.35)**	14.62 (12.59)	**23.09 (5.40)**
h^2 male (SE)	0.02 (0.03)	**0.07 (0.03)**	0.05 (0.04)	**0.19 (0.04)**
CV_A males (SE)	4.75 (3.18)	**7.61 (1.83)**	11.53 (4.97)	**18.68 (2.20)**
V_M (SE)	5.18 (10.30)	4.90 (7.55)	2.26E-07 (5.03E-0.5)	2.76E-05 (2.21E-06)
V_R (SE)	193.39 (13.65)	122.88 (9.79)	222.27 (20.00)	78.35 (6.27)
V_P (SE)	238.22 (8.88)	170.91 (6.33)	235.88 (13.10)	123.05 (4.96)
$COV_{A\ female,male}$ (SE)	**25.73 (6.29)**	**11.75 (4.92)**	-9.24 (11.02)	**11.54 (4.21)**
$r_{A\ female,male}$ (SE)	**1.55 (0.66)**	**0.54 (0.22)**	-0.36 (0.45)	**0.52 (0.18)**

Genetic correlations between fitness in different conditions were always positive (Figure 4) and significant. However, we found significant differences between the different cross-condition correlations (P = 0.05). Correlations were slightly lower when the conditions differed in temperature and humidity (Control–Hot-Dry, Dry-Hot) (Figure 4). We found significant G x E between Control and Dry, Control and Hot-Dry as well as Dry and Hot.

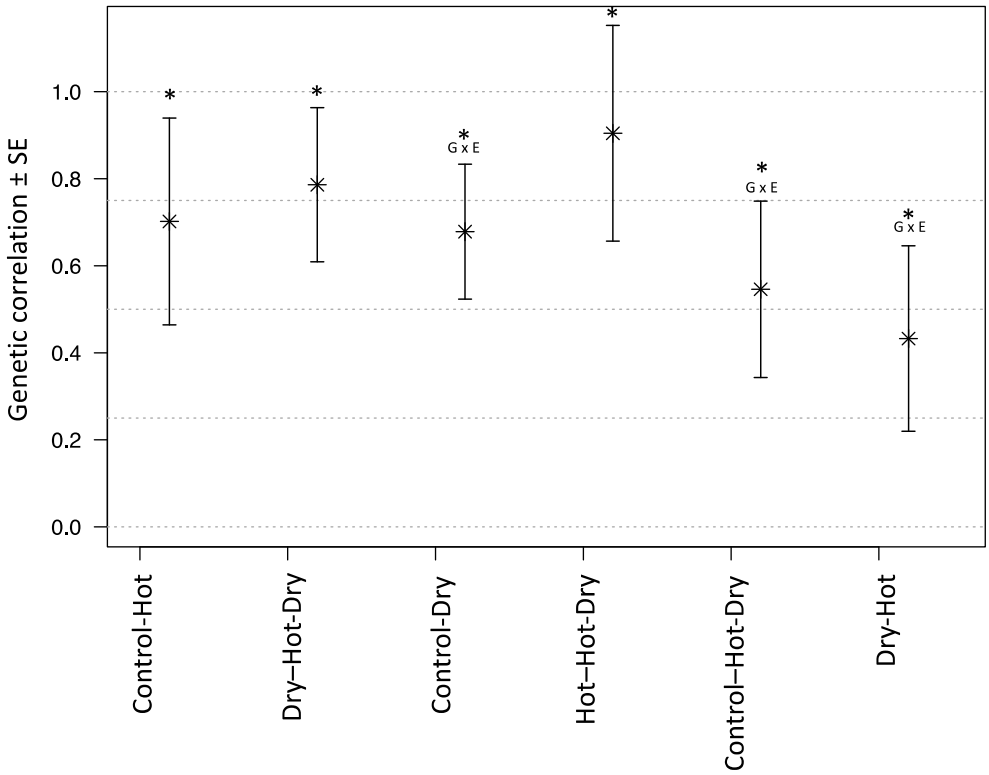


Figure 4: Pairwise cross-environment genetic correlations of fitness in *Tribolium castaneum* estimated using bivariate animal models. * show significant correlations, i.e. significantly different from zero. G x E (= genotype by environment interactions) indicate correlations significantly different from one. Control: 33°C, 70% relative humidity (r.h.); Dry: 33°C, 30% r.h.; Hot: 37°C, 70% r.h.; Hot-Dry: 37°C, 30% r.h.

Genetic correlations in fitness between females and males (Figure 5) were significant in Control (P = 1.02E-06), Dry (P = 1.34E-02), and Hot-Dry (P = 4.47E-04), but not in Hot (P = 0.10). The correlation was highest in Control (Table 1, Figure 5). It decreased in the stress treatments and became even negative in Hot. In all conditions except Control, we found that the correlation was significantly different from one (Dry: P = 1.95E-03; Hot: P = 5.78E-03; Hot-Dry: P = 1.26E-04), indicating G x S, i.e. that genetic basis of fitness is different in the sexes. I_A, the expected evolutionary change as percentage of the mean, was highest in Hot and smallest in Control

(females/males ± SE: Control: 1.59 ± 0.49/0.30 ± 0.29; Dry: 2.05 ± 0.50/0.63 ± 0.27; Hot: 4.32 ± 1.81/1.25 ± 1.14; Hot-Dry: 3.20 ± 1.14/ 3.51 ± 8.1). Based on this, we should observe adaptation in all stress treatments with the largest relative fitness increase in Hot. However, given the negative correlation between male and female additive genetic effects (Figure 5, Table 1), the evolutionary response may be constrained and less than predicted. Taking female – male genetic correlations into account (see Methods) gave us an estimated increase of mean fitness of 3.90% in Dry, 3.89% in Hot, and 10.20% in Hot-Dry per generation. Assuming V_A remains constant, the total change of mean fitness after 19 generations (generation 20) would be 106.81% in Dry, 106.57% in Hot and 532.54% in Hot-Dry.

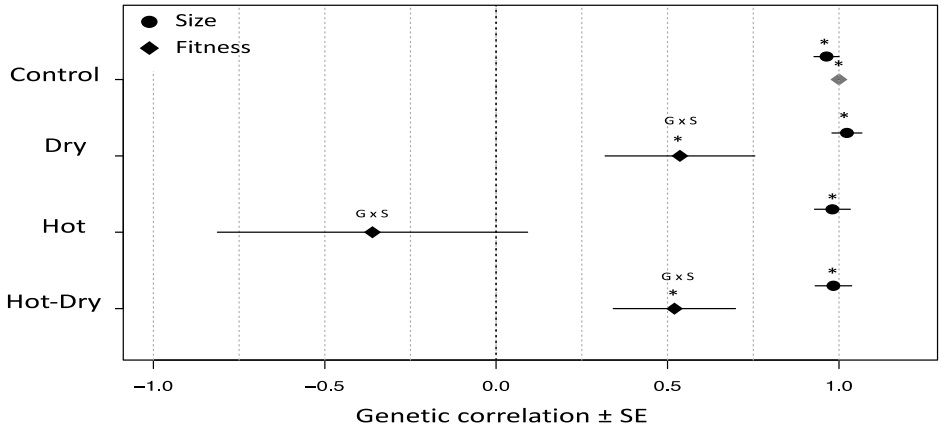


Figure 5: Cross-sex genetic correlations for size and fitness in *Tribolium castaneum* under different environmental conditions. * indicates a significant correlation. G x S (= genotype by sex interaction) indicates that correlation is significantly different from one. Estimate for genetic correlation of fitness in Control was bounded at one and no SE was available. Using models with unconstrained variances and covariances yielded an estimate of 1.55 ± 0.66. V_A of males was small in Control, which impedes precise estimation of cross-sex genetic correlation.

Size

V_A in size was highly significant in both sexes and in all conditions ($P < 0.001$). The results for V_D were less clear. In Control and Hot-Dry estimates for V_D were extremely small and estimates bounded at zero, indicating that

non-additive effects contributed little to observed variation (Figure 6). In Dry and Hot, V_D was not significant, but variance estimates were high when it was included in the model (Table 2, Figure 6). A model without non-additive effects resulted in much higher estimates of V_A in Dry and Hot (Supplemental Table S3). In most cases, P-values for V_D were far from significance (P > 0.2), but for females in Dry and in Hot P-values were lower although still not significant (P = 0.13 and P = 0.10). When we combined males and females and added sex as fixed effect in the model, we obtained a P-value of 0.075 in Hot. Genetic correlations for size across conditions were always positive (Supplemental Figure S7), but no clear pattern emerged. We found significant G x E in female size between Control-Dry (P = 2.78E-03). In male size, G x E was significant between Control-Dry (P = 6.39E-03), Dry-Hot (P = 5.39E-03), Dry–Hot-Dry (P = 0.04), and Hot–Hot-Dry (P = 3.93E-03). Genetic correlations between female and male size were close to one (Table 2, Figure 5) in all conditions, suggesting that size cannot evolve independently in both sexes. I_A of female and male size was not significantly different (P = 0.08) considering all conditions. Although differences in Control seemed to be substantial (Figure 4B-D), they were not significant (h^2: P = 0.26; I_A: P = 0.06). We found that environmental change did not influence I_A and h^2 of female size (P = 0.13; P = 0.14) nor male I_A (P = 0.13), but had a significant effect on male h^2 (P = 0.02).

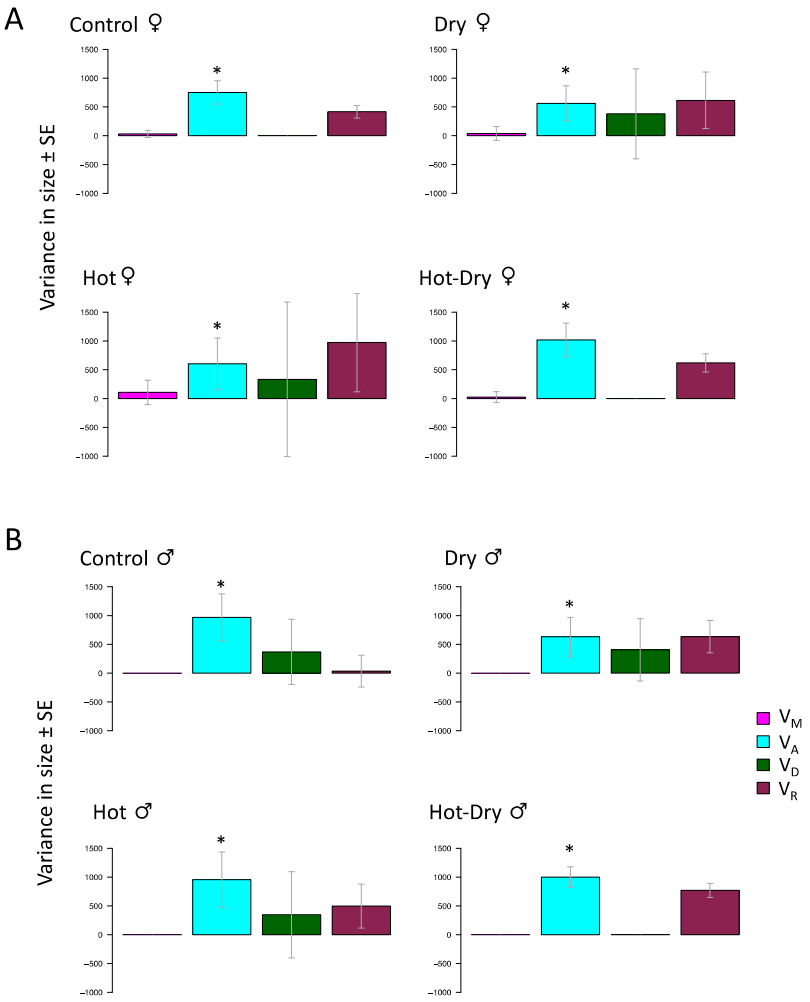


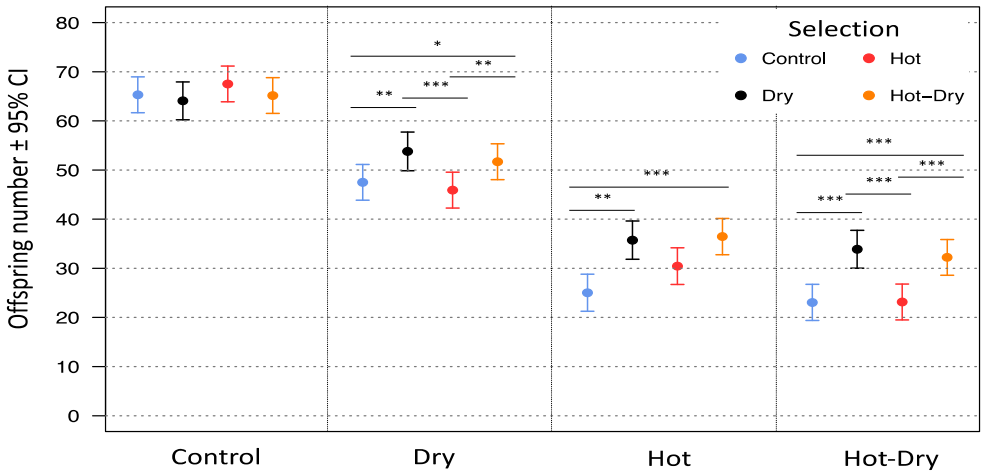
Figure 6: Variance components of size in female **(A)** and male **(B)** flour beetles (*Tribolium castaneum*) under four environmental conditions: V_M: maternal variance, V_A: additive genetic variance, V_D: non-additive genetic variance, V_R: residual variance. * indicates a significant variance component.

Table 2: Body size genetic variances (V_A: additive genetic variance; V_D: dominance variance; V_M: maternal variance; V_R: residual variance), heritability (h^2), coefficient of additive genetic variance (CV_A) in female and male flour beetles and cross-sex additive genetic covariances (COV_A) and correlations (r_A) in different environmental conditions. Estimates for genetic variances were obtained from univariate animal models for each sex separately. For COV_A and r_A bivariate animal models were used. All significant results are in bold.

	Condition							
	Control		Dry		Hot		Hot-Dry	
Females N	1020		1181		837		960	
mean (SE)	1149.13	(1.10)	1148.08	(1.18)	1126.25	(1.57)	1141.96	(1.29)
V_A (SE)	**751.79**	**(201.63)**	**560.98**	**(305.17)**	**603.78**	**(446.35)**	**1017.45**	**(291.52)**
V_D (SE)	6.63E-04	(1.73E-04)	379.526	(782.07)	333.554	(1342.287)	8.670E-04	(2.21E-04)
V_M (SE)	29.61	(63.13)	38.57	(119.09)	107.14	(210.86)	25.58	(96.18)
V_R (SE)	417.11	(108.70)	613.14	(490.95)	973.23	(854.49)	618.81	(157.41)
V_P (SE)	1198.52	(72.27)	1592.19	(82.03)	2017.74	(117.05)	1661.84	(100.31)
h^2 (SE)	**0.63**	**(0.14)**	**0.35**	**(0.18)**	**0.30**	**(0.21)**	**0.61**	**(0.15**
CV_A (SE)	**2.39**	**(0.32)**	**2.06**	**(0.56)**	**2.18**	**(0.81)**	**2.79**	**(0.40)**
Males N	1008		1183		834		962	
mean (SE)	1132.31	(1.18)	1125.47	(1.19)	1103.73	(1.50)	1114.01	(1.34)
V_A (SE)	**969.79**	**(406.67)**	**632.66**	**(335.68)**	**954.83**	**(479.70)**	**999.70**	**(179.66)**
V_D (SE)	369.18	(567.38)	404.65	(539.19)	346.49	(748.68)	4.94E-03	(7.90E-04)
V_M (SE)	6.59E-06	(5.07E-05)	4.88E-04	(2.18E-04)	6.54E-04	(5.00E-04)	3.14E-04	(5.02E-05)
V_R (SE)	35.85	(275.83)	634.64	(282.89)	496.90	(380.49)	769.42	(123.08)
V_P (SE)	1374.81	(293.93)	1672.04	(87.39)	798.22	(118.96)	1769	(98.50)
h^2 (SE)	**0.71**	**(0.25)**	**0.38**	**(0.19)**	**0.53**	**(0.25)**	**0.57**	**(0.08)**
CV_A (SE)	**2.75**	**(0.58)**	**2.23**	**(0.59)**	**2.80**	**(0.70)**	**2.84**	**(0.26)**
$COV_{A\ female,male}$ (SE)	**955.25**	**(120.93)**	**1002.06**	**(128.10)**	**1140.11**	**(175.33)**	**976.20**	**(137.10)**
$r_{A\ female,male}$ (SE)	**0.96**	**(0.04)**	**1.02**	**(0.04)**	**0.98**	**(0.05)**	**0.98**	**(0.05)**

Adaptation after 20 generations and asymmetric correlated responses

We found significant adaptation to all conditions after 20 generations, shown by significant effects of selection regime on offspring number (Dry: $F_{1,16} = 10.43$, $p = 0.005$; Hot: $F_{1,18} = 4.80$, $p = 0.042$; Hot-Dry: $F_{1,18} = 14.78$, $p = 0.001$, Table S2). In all treatments, the native selection lines produced significantly more offspring than non-adapted control lines (Figure 7). The largest difference and most significant fitness increase was observed in the most stressful condition Hot-Dry. In contrast to the three stress treatments, we did not find any differences between lines from different selection regimes in the ancestral control condition (Figure 7), showing that adaptation to stress treatments did not come at a cost of reduced fitness in control conditions. Interestingly, drought selection resulted in higher heat resistance. Dry lines showed a significantly higher offspring number in Hot and Hot-Dry (estimated increase in offspring number relative to Control lines: 12.10 ± 2.23, p = 3.64E-07 in Hot; 12.18 ± 2.18, p = 4.20E-07 in Hot-Dry). In Hot, they performed even better than the native Hot-lines (mean offspring number in Hot [95% CI]: Control-lines: 24.59 [20.71, 28.47]; Hot-lines: 29.94 [26.11, 33.77]; Dry-lines: 35.69 [31.77, 39.61]; Hot-Dry-lines: 36.91 [33.14, 40.68]). We did not observe such a correlated response in the Hot-lines (Figure 7). Their offspring number in Dry was not different from those of Control-Lines. Lines in Hot-Dry that adapted to a combination of heat and drought showed an increased fitness in both single stressor treatments (in Dry: 4.45 ± 2.10, p = 0.04; in Hot: 11.66 ± 2.17, p = 1.20E-04). Despite high genetic correlation in fitness between Hot and Hot-Dry estimated in the first generation, we found no correlated response of Hot-lines to Hot-Dry conditions. Their offspring number was not different from non-adapted Control-Lines (Hot-lines: 23.14 [19.43, 26.84 CI]; Control-lines: 23.03 [19.30, 26.76 CI]).



414

415

416 **Figure 7**: Mean offspring number per female of lines from different selection regimes under different environmental
417 conditions. Significant differences between selection regimes within the same condition (Tukey's HSD post hoc test, p-
418 values adjusted for multiple comparisons) are indicated by *P<0.05, **P<0.01, ***P<0.001.

419

420 The observed changes in fitness after experimental evolution were much lower than the predicted total
421 changes. Observed fitness increases relative to Control-lines were: Dry: 16.43 ± 3.24 %, Hot: 22.37 ± 4.34 %
422 and 46.29 ± 2.17 % in Hot-Dry. Nonetheless, in agreement with our predictions we observed the strongest
423 increase in Hot-Dry and similar increases in Dry and Hot.

424

425

426

427

428

Discussion

Consequences of environmental change on persistence of populations strongly depend on their ability to adapt either by plastic or genetic changes. The stress conditions we applied, had a strong impact on fitness. Offspring number in Hot-Dry was reduced to ca. 50% of control level thus clearly showing that plasticity was not able to fully compensate for the negative effects of environmental change. We found significant V_A in fitness in all conditions indicating potential to adapt by genetic changes with no evidence that environmental changes lowered the adaptive potential. In contrary, I_A of female and male fitness increased in the stress treatments, which should facilitate adaptation. Accordingly, we found significant adaptation to all conditions after 20 generations of experimental evolution. We also did not find strong evidence of genetic constraints on adaptation between stressors or sexes as most genetic correlations were positive (but not in Hot, see Figure 5). Taking account of both male and female variances and covariance in fitness allowed us to make better predictions of fitness increase. All in all, our study shows that although it is not possible to predict the precise fitness increase over this period of 20 generations, we can make qualitative predictions about adaptation based on V_A estimates. We can examine whether a population is likely to adapt and make relative comparisons between stress treatments, i.e. identify those where we expect the largest relative increase in fitness. We thus show that existing quantitative genetic tools are informative over time scales beyond single generation responses. However, reliability of predictions may require a fully integrative approach including potential genetic covariances between the sexes as well as a careful choice of the traits used as fitness estimate.

Evolutionary potential and adaptation

V_A of female fitness was higher than V_A of male fitness in all conditions except Hot-Dry. Egg production is likely to be costly and reduced when females are stressed due to a resource allocation trade-off. It is therefore not surprising that variation in reproductive output could be mainly explained by genetic differences among females, whereas genetic variation among males had only a minor influence. However, this changed when humidity was reduced. In Dry conditions, V_A for males became significant (but was still much smaller compared

to females) and in Hot-Dry, V_A in both sexes were similar. It was observed that the condition of *Tribolium* males could have a significant influence on the reproductive output of their mating partner. Starvation of males reduced insemination success (Lewis et al. 2012), the number of eggs laid, and the proportion of unfertilized eggs (Sbilordo et al. 2011). It seems plausible that desiccation might have similar effects. Producing ejaculate may be costly and decreased under dry conditions to reduce water loss. Genetic differences in male drought resistance would then translate into observed offspring variation and result in a higher male influence on female reproductive output. Although male V_A was small in ancestral conditions, evolution of males seemed to play an important role in adapting to Hot-Dry. In contrast to the between-sex differences in fitness V_A, we did not find clear differences in V_A between sexes for body size. The largest difference was in Control (Figure 2), with a lower female V_A. It is likely that female size had been under strong positive selection since it is associated with fecundity (Honěk 1993), which might explain the observed reduction. Additionally, while we observed an increase in V_A for male fitness in the Dry treatment, V_A for male size was reduced in Dry. A drought effect was thus also detectable for size: When we explored the effects of temperature and humidity on size, we found a significant (although small) sex-treatment-interaction indicating that male size was more sensitive to drought than female size. G x E for male size occurred when conditions differed in humidity.

Positive cross-condition genetic correlations indicated an absence of an evolutionary trade-off between drought and heat adaptation. Consistently, we did not find any selection line performing worse than control lines in any condition after experimental evolution. It was previously shown that the treatments induce substantial changes in gene expression (Koch and Guillaume 2020). It might then be that selection in the stress treatments is on genes that are not expressed in control and thus allele frequency changes do not sufficiently affect fitness under control conditions. Because positive genetic correlations indicate a similar genetic basis for fitness under different conditions, adapting to one condition should result in a correlated response and increased fitness in other conditions. Accordingly, we found that drought adaptation improved heat resistance (Figure 7). However, this correlated response was asymmetric since we did not observe an effect of heat adaptation on drought resistance. A possible explanation is that selection in dry and hot conditions shifted

481 allele frequencies of genes with different pleiotropic effects (Bohren et al. 1966). Alternatively, the pleiotropic
482 degree of a gene might be environment dependent (Barrett et al. 2009). Interestingly, we found that genetic
483 correlations in generation one were lower when the conditions differed in both temperature and humidity,
484 which is the case in Control–Hot-Dry and Dry-Hot (Figure 4) suggesting that the genetic basis for adaptation to
485 heat and drought might be slightly different.
486
487 We found the highest positive cross-sex genetic correlation in fitness in Control with lower estimates in the
488 treatments and a negative correlation in Hot. Several studies in the wild (Brommer et al. 2007; Foerster et al.
489 2007; Mokkonen et al. 2011) as well as in laboratory populations (Delcourt et al. 2009; Punzalan et al. 2014)
490 reported negative genetic correlations between female and male fitness. It is not well understood how cross-
491 sex correlations are influenced by the environment and contrasting predictions have been made. For instance,
492 it was argued that in a population far from its optimum after an environmental shift, females and males might
493 experience similar directional selection leading to high a positive correlation in their fitness (Berger et al. 2014;
494 Connallon and Hall 2016; Wolak et al. 2018). In contrast, genetic covariances between sexes have been
495 predicted to be less negative in ancestral conditions (Delcourt et al. 2009), because selection on the long-term
496 should favour alleles providing fitness benefits to both (Collet et al. 2016). This last prediction is consistent
497 with our observations.
498
499 Taking the male V_A and the covariances between female and male fitness into account led us to predict the
500 highest relative increase in fitness in Hot-Dry and smaller but similar increases in Dry and Hot. Ignoring male
501 and cross-sex effects when studying the adaptive capacity of a population can thus lead to misleading
502 predictions. The results of the fitness assay after 20 generations mainly matched our predictions but increases
503 in fitness were much lower than our estimated upper limit. However, our predictions are based on the
504 assumption of constant V_A and consequently exponentially increasing fitness (Falconer and MacKay 1996).
505 Trade-offs between different fitness components are expected to occur because physiological limits, e.g. for

egg laying rate, exist and should prevent an infinite fitness increase. It is also important to note that the conditions during the fitness assay were not exactly the same as during evolution. To measure the offspring number per female. each female was kept individually in egg-laying tubes. In contrast, all individuals of the same line were kept in one vial during experimental evolution. Effects of density and competition that might have influenced adaptation were not captured in our fitness assay. Our results after 19 generations might be further influenced by evolution of control lines that were used as reference representing the ancestral non-evolved stage. Significant V_A under control conditions in addition to high and positive genetic correlations between fitness in control and stress conditions suggest a fitness increase in control lines over time and correlated responses in the treatments. This might have led to an underestimation of the true fitness increase of adapted selection lines. Interestingly, we found no difference between lines from different selection regimes in Control. A reason might be that genes responsible for fitness in the treatments are not expressed under control conditions. Evolutionary changes in these genes that occurred during adaptation to the treatments would not show any effect on fitness under control conditions leading to cryptic genetic variation. In control lines these genes were not exposed to selection and consequently control lines showed a lower fitness compared to adapted selection line when exposed to the treatments.

Variance components of fitness and body size

Both traits differed markedly in the proportion of different variance components and in the environmental effect on genetic variances. h^2 was much higher in size than in fitness. This is a common finding in many studies (Roff and Mousseau 1987; Kruuk et al. 2000; McCleery et al. 2004; Teplitsky et al. 2009). It had been initially interpreted as evidence for strong selection depleting V_A in fitness related traits. However, a high V_A can be concealed in h^2 by a simultaneous increase of the total variance (Houle 1992; Merilä and Sheldon 1999; Hansen et al. 2011; Wheelwright et al. 2014). Given the highly polygenic nature of fitness, it was even argued that fitness might show higher V_A (Merilä and Sheldon 1999), since it represents a larger mutational target. Furthermore, fitness is a composite character with a high number of morphological, physiological and behavioral traits contributing to it, each of them with some underlying genetic variance. However, each

532 contributing trait may increase the influence of environmental variation, leading to a higher total variance in
533 fitness (Price and Schulter 1991) and thus a lower h^2. According to those previous considerations, we found
534 much higher estimates of mean-scaled V_A (CV_A, I_A) in offspring number than in body size and a lower h^2 in
535 fitness. The lower h^2 was mainly due to a higher proportion of environmental variance (V_R) when compared to
536 body size. Contrary to body size, we could not directly estimate V_D for fitness because we could not
537 disentangle V_D from V_M and common environmental effects in the F1 with our half-sib/full-sib breeding design.
538 However, if V_D were present, then it should be included in V_M, because it contributes to full-sib resemblance
539 via a shared mother. Comparative studies investigating the relative amount of V_A and V_D suggested that
540 proportion of V_D can be substantial and of the same magnitude as V_A (Crnokrak and Roff 1995; Wolak and
541 Keller 2014). An increased proportion of V_D is expected in populations under strong selection, since V_D is not
542 affected by natural selection (Crnokrak and Roff 1995; Merilä and Sheldon 1999; Roff and Emerson 2006;
543 Sztepanacz and Blows 2015), or with increased inbreeding (Falconer and Makay, 1996). In our experiment, V_M
544 for fitness was much lower than V_A, even close to zero in Hot and Hot-Dry (Figure 3), suggesting that V_D
545 contributed much less to total genetic variance than V_A.
546
547 We could directly estimate V_D for body size in the F2 cross and can thus provide data for this rarely estimated
548 variance component. Although our V_D estimates for size were always associated with large SE, they suggest
549 that V_D is present and environment dependent. The highest estimates of V_D were found in Dry and Hot for both
550 male and female size, while it remained close to zero in Control and Hot-Dry. This environmental dependence
551 of V_D has not been studied in detail before, although the environmental dependence of inbreeding depression
552 was previously described (Bijlsma et al. 1999; Armbruster and Reed 2005; Fox and Reed 2011). Both V_D and
553 inbreeding depression are expected to increase with inbreeding, for instance in shrinking populations when
554 allele and genotype frequencies change because of increased drift (Falconer and Makay, 1996). This could have
555 important implications since environmental changes may emphasize the effects of inbreeding on survival and
556 thus on population size too, directly affecting genetic variances.

The large uncertainty around our estimates of V_D may come from the double-first cousin (DFC) breeding design used in the F2, despite the large sample size we had (827 family pairs with DFC offspring). Although the DFC design was proposed to estimate V_D (Fairbairn and Roff 2006), it may have limited statistical power because the probability that DFC share alleles identical by descent at a given locus is only 6.25 %, resulting in large SE, in contrast to full-sibs where this probability is 25%. Therefore, comparing maternal half-sibs and full-sibs might be a much more powerful approach to estimate V_D, while it allows us to disentangle non-additive from maternal effects at the same time .

Although we could estimate genetic correlations between fitness and size with our data set (Supplemental Table S4), it was not possible to get unbiased estimates of those correlations. With our design, we could not rule out a strong confounding effect of population density because body size was measured in the F2, the offspring of beetles used for the fitness assay (Supplement Figure S1). Consequently, offspring of females with a high fitness (i.e. high offspring number) grew at a higher density since we used identical tubes with the same amount of flour for all females in the fitness assay.

Conclusions

Our study showed an increased adaptive potential in stressful conditions and a corresponding adaptation to those conditions after 20 generations. Although precise predictions of relative increase in fitness were not possible over this time period, we could make correct qualitative predictions. We expected and observed the highest relative fitness increase in the most stressful hot-dry condition and similar increases in single stress treatments. The apparently high adaptive potential of female beetles in Hot was limited by a negative genetic correlation with male fitness. We further found that genetic effects of males on fitness can be large and can increase the adaptive potential. Comparing genetic variances of fitness and size showed that they differed in their variance composition and in their cross-sex and cross-environment genetic covariances. Environmental effects on genetic variances were also not consistent between the two traits. We thus advise caution if studies

583 interested in fitness V_A and the adaptive capacity of a population use body size as a proxy. Overall, we found
584 that genetic variance in fitness is a key estimate of a population's adaptive capacity for time scales over 20
585 generations. As such, it may help predict the adaptive response of populations exposed to new environmental
586 conditions and help identify the populations most at risk of extinction. However, the reliability of such
587 predictions will depend on the fitness estimate chosen and on the full integration of the multifaceted aspects
588 of adaptation. Inclusion of genetic covariances between female and male fitness and genotype by environment
589 interactions is thus important.
590
591 **Author contributions:** FG and ELK designed the experiment. ELK and SHS conducted the experiment (crossing
592 and fitness assay). SHS performed size measurements. ELK analysed the data. FG and ELK wrote the
593 manuscript. SHS contributed and commented to manuscript.
594
595 **Acknowledgements**: This work was supported by the Swiss National Science Foundation, grants
596 PP00P3_1144846 and PP00P3_176965 to FG. We would like to thank all people helping with fitness
597 measurements: Julian Bauer, Valérian Zeender, Maria Domenica Moccia, Sara Meier, Cilgia Lippuner, Tim
598 Schoch, Tim Emmenegger, Elke Karaus.
599

600 ## References

601 Armbruster, P., and D. H. Reed. 2005. Inbreeding depression in benign and stressful environments. Heredity
602 95:235–242.
603 Arnold, P. A., A. B. Nicotra, and L. E. B. Kruuk. 2019. Sparse evidence for selection on phenotypic plasticity in
604 response to temperature. Philos. Trans. R. Soc. B Biol. Sci. 374.
605 Barrett, R. D. H., S. M. Rogers, and D. Schluter. 2009. Environment specific pleiotropy facilitates divergence at
606 the ectodysplasin locus in threespine stickleback. Evolution 63:2831–2837.
607 Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat.
608 Softw. 67.

609 Bell, G., and A. Gonzalez. 2009. Evolutionary rescue can prevent extinction following environmental change.
610 Ecol. Lett. 12:942–948.
611 Berger, D., K. Grieshop, M. I. Lind, J. Goenaga, A. A. Maklakov, and G. Arnqvist. 2014. Intralocus sexual conflict
612 and environmental stress. Evolution 68:2184–2196.
613 Bijlsma, R., J. Bundgaard, and W. F. Van Putten. 1999. Environmental dependence of inbreeding depression
614 and purging in Drosophila melanogaster. J. Evol. Biol. 12:1125–1137.
615 Bijma, P., and M. J. Wade. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on
616 response to genetic selection. J. Evol. Biol. 21:1175–1188.
617 Bohren, B. B., W. G. Hill, and A. Robertson. 1966. Some observations on asymmetrical correlated responses to
618 selection. Genet. Res. 7:44–57.
619 Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol. Evol. 24:280–288.
620 Brommer, J. E., M. Kirkpatrick, A. Qvamström, and L. Gustafsson. 2007. The intersexual genetic correlation for
621 lifetime fitness in the wild and its implications for sexual selection. PLoS One 2:e744. Public Library of
622 Science.
623 Bubliy, O. A., T. N. Kristensen, V. Kellermann, and V. Loeschcke. 2012. Plastic responses to four environmental
624 stresses and cross-resistance in a laboratory population of *Drosophila melanogaster*. Funct. Ecol. 26:245–
625 253.
626 Bubliy, O. A., and V. Loeschcke. 2005. Correlated responses to selection for stress resistance and longevity in a
627 laboratory population of *Drosophila melanogaster*. J. Evol. Biol. 18:789–803.
628 Butler, D. G., B. R. Cullis, A. R. Gilmour, and B. J. Gogel. 2009. Analysis of mixed models for S language
629 environments: ASReml-R reference manual. The State of Queensland, Department of Primary Industries
630 and Fisheries, Brisbane.
631 Calsbeek, R., M. C. Duryea, D. Goedert, P. Bergeron, and R. M. Cox. 2015. Intralocus sexual conflict, adaptive
632 sex allocation, and the heritability of fitness. J. Evol. Biol. 28:1975–1985.
633 Charmantier, A., and D. Garant. 2005. Environmental quality and evolutionary potential: lessons from wild
634 populations. Proc. R. Soc. B Biol. Sci. 272:1415–1425.
635 Charmantier, A., D. Garant, and L. E. B. Kruuk (eds). 2014. Quantitative genetics in the wild. Oxford University

636 Press, Oxford.

637 Charmantier, A., R. H. McCleery, L. R. Cole, C. Perrins, L. E. B. Kruuk, and B. C. Sheldon. 2008. Adaptive
638 phenotypic plasticity in response to climate change in a wild bird population. Science 320:800–803.

639 Collet, J. M., S. Fuentes, J. Hesketh, M. S. Hill, P. Innocenti, E. H. Morrow, K. Fowler, and M. Reuter. 2016. Rapid
640 evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. Evolution 70:781–
641 795.

642 Connallon, T., and M. D. Hall. 2016. Genetic correlations and sex-specific adaptation in changing environments.
643 Evolution 70:2186–2198.

644 Crnokrak, P., and D. A. Roff. 1995. Dominance variance: Associations with selection and fitness. Heredity
645 75:530–540. Nature Publishing Group.

646 de Jong, G. 2005. Evolution of phenotypic plasticity: patterns of plasticity and the emergence of ecotypes. New
647 Phytol. 166:101–118.

648 de Villemereuil, P. 2018. Quantitative genetic methods depending on the nature of the phenotypic trait. Ann.
649 N. Y. Acad. Sci. 1422:29–47.

650 Delcourt, M., M. W. Blows, and H. D. Rundle. 2009. Sexually antagonistic genetic variance for fitness in an
651 ancestral and a novel environment. Proc. R. Soc. B Biol. Sci. 276:2009–2014.

652 DeWitt, T. J., A. Sih, and D. S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends Ecol. Evol. 13:77–
653 81.

654 Duputié, A., A. Rutschmann, O. Ronce, and I. Chuine. 2015. Phenological plasticity will not help all species
655 adapt to climate change. Glob. Chang. Biol. 21:3062–3073.

656 Etterson, J. R., and R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science
657 294:151–154.

658 Fairbairn, D. J., and D. A. Roff. 2006. The quantitative genetics of sexual dimorphism: Assessing the importance
659 of sex-linkage. Heredity 97:319–328.

660 Falconer, D. S., and T. F. C. MacKay. 1996. Introduction to Quantitative Genetics. 4th Editio. Longman Group,
661 Essex.

662 Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, UK.

663 Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually
664 antagonistic genetic variation for fitness in red deer. Nature 447:1107–1110.
665 Fox, C. W., and D. H. Reed. 2011. Inbreeding depression increases with environmental stress: An experimental
666 study and meta-analysis. Evolution 65:246–258.
667 Fry, J. D. 2008. Genotype-environment interaction for total fitness in Drosophila. J. Genet. 87:355–362.
668 Fugère, V., and A. P. Hendry. 2018. Human influences on the strength of phenotypic selection. Proc. Natl. Acad.
669 Sci. U. S. A. 115:10070–10075.
670 Garcia-Gonzalez, F., L. W. Simmons, J. L. Tomkins, J. S. Kotiaho, and J. P. Evans. 2012. Comparing evolvabilities:
671 Common errors surrounding the calculation and use of coefficients of additive genetic variation.
672 Evolution 66:2341–2349.
673 Gardner, M. P., K. Fowler, N. H. Barton, and L. Partridge. 2005. Genetic variation for total fitness in Drosophila
674 melanogaster: Complex yet replicable patterns. Genetics 169:1553–1571. Genetics.
675 Gienapp, P., S. Fior, F. Guillaume, J. R. Lasky, V. L. Sork, and K. Csilléry. 2017. Genomic Quantitative Genetics to
676 Study Evolution in the Wild. Trends Ecol. Evol. 32:897–908. Elsevier Ltd.
677 Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN
678 International Ltd, Hemel Hempstead, UK.
679 Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not Evolvability. Evol. Biol. 38:258–277.
680 Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The Contemporary Evolution of Fitness. Annu.
681 Rev. Ecol. Evol. Syst. 49:457–476.
682 Hereford, J. 2009. A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. Am. Nat. 173:579–588.
683 Hoffmann, A. A., and J. Merilä. 1999. Heritable variation and evolution under favourable and unfavourable
684 conditions. Trends Ecol. Evol. 14:96–101.
685 Honěk, A. 1993. Intraspecific Variation in Body Size and Fecundity in Insects : A General Relationship. Nord.
686 Soc. Oikos 66:483–492.
687 Houle, D. 1992. Comparing Evolvability and Variability. Genetics 130:195–204.
688 Husby, A., M. E. Visser, and L. E. B. Kruuk. 2011. Speeding up microevolution: The effects of increasing
689 temperature on selection and genetic variance in a wild bird population. PLoS Biol. 9:e1000585.

690 Kellermann, V. M., B. van Heerwaarden, A. A. Hoffmann, and C. M. Sgrò. 2006. Very Low Additive Genetic
691 Variance and Evolutionary Potential in Multiple. Evolution 60:1104–1108.
692 Kirkpatrick, M. 2009. Patterns of quantitative genetic variation in multiple dimensions. Genetica 136:271–284.
693 Koch, E. L., and F. Guillaume. 2020. Additive and mostly adaptive plastic responses of gene expression to
694 multiple stress in Tribolium castaneum. PLoS Genet. 16:e1008768.
695 Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the “animal model.” Philos.
696 Trans. R. Soc. B Biol. Sci. 359:873–890. The Royal Society.
697 Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E. Guinness. 2000.
698 Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci. 97:698–703.
699 Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects
700 Models. J. Stat. Softw. 82:1–26.
701 Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry.
702 Evolution 33:402–416. Wiley/Blackwell (10.1111).
703 Lenth, R. V. 2016. Least-Squares Means: The R Package lsmeans. J. Stat. Softw. 69:1–33.
704 Lewis, S. M., N. Tigreros, T. Fedina, and Q. L. Ming. 2012. Genetic and nutritional effects on male traits and
705 reproductive performance in *Tribolium* flour beetles. J. Evol. Biol. 25:438–451.
706 Long, T. A. F., P. M. Miller, A. D. Stewart, and W. R. Rice. 2009. Estimating the heritability of female lifetime
707 fecundity in a locally adapted Drosophila melanogaster population. J. Evol. Biol. 22:637–643.
708 McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins. 2004. Components of
709 Variance Underlying Fitness in a Natural Population of the Great Tit Parus major. Am. Nat. 164:E62–E72.
710 McFarlane, S. E., J. C. Gorrell, D. W. Coltman, M. M. Humphries, S. Boutin, and A. G. Mcadam. 2014. Very low
711 levels of direct additive genetic variance in fitness and fitness components in a red squirrel population.
712 Ecol. Evol. 4:1729–1738.
713 Merilä, J., and B. C. Sheldon. 1999. Genetic architecture of fitness and nonfitness traits: empirical patterns and
714 development of ideas. Heredity 83:103–109.
715 Merilä, J., and B. C. Sheldon. 2000. Lifetime Reproductive Success and Heritability in Nature. Am. Nat. 155:301–
716 310.

717 Milutinović, B., C. Stolpe, R. Peuß, S. A. O. Armitage, and J. Kurtz. 2013. The Red Flour Beetle as a Model for
718 Bacterial Oral Infections. PLoS One 8.
719 Mokkonen, M., H. Kokko, E. Koskela, J. Lehtonen, T. Mappes, H. Martiskainen, and S. C. Mills. 2011. Negative
720 frequency-dependent selection of sexually antagonistic alleles in Myodes glareolus. Science 334:972–
721 974. American Association for the Advancement of Science.
722 Morrissey, M. B., and T. Bonnet. 2019. Analogues of the fundamental and secondary theorems of selection,
723 assuming a log-normal distribution of expected fitness. J. Hered. 110:396–402. Oxford Academic.
724 Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity
725 59:181–197.
726 Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. Annu. Rev. Ecol. Evol.
727 Syst. 37:637–669.
728 Perrier, C., B. Delahaie, and A. Charmantier. 2018. Heritability estimates from genomewide relatedness
729 matrices in wild populations: Application to a passerine, using a small sample size. Mol. Ecol. Resour.
730 18:838–853.
731 Pfennig, D. W., M. A. Wund, E. C. Snell-Rood, T. Cruickshank, C. D. Schlichting, and A. P. Moczek. 2010.
732 Phenotypic plasticity's impacts on diversification and speciation. Trends Ecol. Evol. 25:459–467.
733 Poissant, J., A. J. Wilson, and D. W. Coltman. 2010. Sex-specific genetic variance and the evolution of sexual
734 dimorphism: A systematic review of cross-sex genetic correlations. Evolution 64:97–107.
735 Price, G. R. 1972. Fisher's 'fundamental theorem' made clear. Ann. Hum. Genet. 36:129–140.
736 Price, T., and D. Schulter. 1991. On the Low Heritability of Life-History Traits. Evolution 45:853–861.
737 Punzalan, D., M. Delcourt, and H. D. Rundle. 2014. Comparing the intersex genetic correlation for fitness across
738 novel environments in the fruit fly, *Drosophila serrata*. Heredity 112:143–148.
739 R Core Team. 2017. R Core Team (2017). R: A language and environment for statistical computing. R Found.
740 Stat. Comput. Vienna, Austria. URL http//www.R-project.org/.
741 Roff. 2000. Trade-offs between growth and reproduction: an analysis of the quantitative genetic evidence. J.
742 Evol. Biol. 13:434–445.).
743 Roff, D. A., and K. Emerson. 2006. Epistasis and dominance: evidence for differential effects in life-history

versus morphological traits. Evolution 60:1981.

Roff, D. A., and T. A. Mousseau. 1987. Quantitative genetics and fitness: lessons from Drosophila. Heredity 58:103–118.

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

Sbilordo, S. H., V. M. Grazer, M. Demont, and O. Y. Martin. 2011. Impacts of starvation on male reproductive success in *Tribolium castaneum*. Evol. Ecol. Res. 13:347–359.

Schluter, D., T. D. Price, and L. Rowe. 1991. Conflicting selection pressures and life history trade-offs.

Sgrò, C. M., and A. A. Hoffmann. 1998. Effects of stress combinations on the expression of additive genetic variation for fecundity in *Drosophila melanogaster*. Genet. Res. 72:13–18.

Sgrò, C. M., and A. A. Hoffmann. 2004. Genetic correlations, tradeoffs and environmental variation. Heredity 93:241–248.

Shaw, R. G., and J. R. Etterson. 2012. Rapid climate change and the rate of adaptation: Insight from experimental quantitative genetics.

Shaw, R. G., and F. H. Shaw. 2014. Quantitative genetic study of the adaptive process. Heredity 112:13–20.

Sikkink, K. L., R. M. Reynolds, W. A. Cresko, and P. C. Phillips. 2015. Environmentally induced changes in correlated responses to selection reveal variable pleiotropy across a complex genetic network. Evolution 69:1128–1142.

Simons, A. M., and D. A. Roff. 1996. The Effect of a Variable Environment on the Genetic Correlation Structure in a Field Cricket. Evolution 50:267–275.

Snell-Rood, E. C., J. D. Van Dyken, T. Cruickshank, M. J. Wade, and A. P. Moczek. 2010. Toward a population genetic framework of developmental evolution: The costs, limits, and consequences of phenotypic plasticity.

Sokolova, I. M., M. Frederich, R. Bagwe, G. Lannig, and A. A. Sukhotin. 2012. Energy homeostasis as an integrative tool for assessing limits of environmental stress tolerance in aquatic invertebrates. Mar. Environ. Res. 79:1–15.

Sztepanacz, J. L., and M. W. Blows. 2015. Dominance genetic variance for traits under directional selection in

Drosophila serrata. Genetics 200:371–384.

Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in a wild bird population. Evolution 63:716–726.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. N. Erasmus, M. Ferreira De Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. Van Jaarsveld, G. F. Midgley, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from climate change. 427:145–148.

Valladares, F., E. Gianoli, and J. M. Gómez. 2007. Ecological limits to plant phenotypic plasticity. New Phytol. 176:749–763.

Waldvogel, A.-M., B. Feldmeyer, G. Rolshausen, M. Exposito-Alonso, C. Rellstab, R. Kofler, T. Mock, K. Schmid, I. Schmitt, T. Bataillon, O. Savolainen, A. Bergland, T. Flatt, F. Guillaume, and M. Pfenninger. 2020. Evolutionary genomics can improve prediction of species' responses to climate change. Evol. Lett. 4:4–18.

Wheelwright, N. T., L. F. Keller, and E. Postma. 2014. The effect of trait type and strength of selection on heritability and evolvability in an island bird population. Evolution 68:3325–3336.

Wickman, P.-O., and B. Karlsson. 1989. Abdomen Size, Body Size and the Reproductive Effort of Insects. Oikos 56:209.

Wilson, A. J., J. M. Pemberton, J. G. Pilkington, D. W. Coltman, D. V Mifsud, T. H. Clutton-Brock, and L. E. B. Kruuk. 2006. Environmental coupling of selection and heritability limits evolution. PLoS Biol. 4:1270–1275.

Wilson, A. J., D. Réale, M. N. Clements, M. M. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79:13–26.

Wolak, M. E. 2012. Nadiv: An R package to create relatedness matrices for estimating non-additive genetic variances in animal models. Methods Ecol. Evol. 3:792–796.

Wolak, M. E., P. Arcese, L. F. Keller, P. Nietlisbach, and J. M. Reid. 2018. Sex-specific additive genetic variances and correlations for fitness in a song sparrow (*Melospiza melodia*) population subject to natural immigration and inbreeding. Evolution 72:2057–2075.

798 Wolak, M. E., and L. F. Keller. 2014. Dominance genetic variance and inbreeding in natural populations. Pp.
799 104–127 *in* A. Charmantier and L. E. B. Kruuk, eds. Quantitative Genetics in the Wild. Ox, Oxford, UK.
800 Wood, C. W., and E. D. Brodie. 2015. Environmental effects on the structure of the G-matrix. Evolution
801 69:2927–2940.
802 Wood, C. W., and E. D. Brodie. 2016. Evolutionary response when selection and genetic variation covary across
803 environments. Ecol. Lett. 19:1189–1200.