Evolution of phenotypic variance provides insights into the genetic basis of adaption

Author list:

Wei-Yun Lai[1,2], Viola Nolte[1], Ana Marija Jakšić[1,2,3] and Christian Schlötterer[1*]

[1]Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria.

[2]Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Vienna, Austria.

[3]Current affiliation: École polytechnique fédérale de Lausanne, Lausanne, Switzerland

[*]Correspondence: christian.schloetterer@vetmeduni.ac.at; Tel.: +43-1-25077-4300.

Keywords:

Phenotypic variance, temperature adaptation, *Drosophila simulans*, experimental evolution

Abstract

Most traits are polygenic and the contributing loci can be identified by GWAS. Their adaptive architecture is, however, poorly characterized. Here, we propose a new approach to study the adaptive architecture, which does not depend on genomic data. Relying on experimental evolution we measure the phenotypic variance in replicated populations during adaptation to a new environment. Extensive computer simulations show that the evolution of phenotypic variance in a replicated experimental evolution setting is a powerful approach to distinguish between oligogenic and polygenic adaptive architectures. We apply this new method to gene expression variance in male *Drosophila simulans* before and after 100 generations of adaptation to a novel hot environment. The variance change in gene expression was indistinguishable for genes with and without a significant change in mean expression after 100 generations of evolution. We conclude that adaptive gene expression evolution is best explained by a highly polygenic adaptive architecture. We propose that the evolution of phenotypic variance provides a powerful approach to characterize the adaptive architecture, in particular when combined with genomic data.

Introduction

It is widely accepted that most complex traits have a polygenic or even infinitesimal basis (Ayroles et al., 2009; Boyle et al., 2017; Liu et al., 2019). Nevertheless, it is difficult to predict which of these loci are responding to selection when a population is exposed to a new selection regime. If pleiotropic constraints are strong, only a small subset of the genes constituting are free to respond to respond to selection. Hence, the genetic basis of the adaptive response of a complex trait (i.e. adaptive architecture (Barghi et al., 2020)) may differ substantially from the genetic architecture. Since even for large phenotypic changes the genetic basis of an adaptive response is difficult to study when more than a handful of genes are contributing, we introduce a new approach to study the complexity of the adaptive architecture. Rather than aiming to map the contributing loci, we propose to study the evolution of phenotypic variance in an experimental evolution framework.

The phenotypic variance of a quantitative trait is a key determinant for its response to selection. It can be decomposed into genetic and environmental components (Falconer and Mackay, 1963). Over the past years, mathematical models have been developed which describe the expected genetic variance of a quantitative trait under selection and its maintenance in evolving populations (Bulmer, 1972; Chevalet, 1994; Kimura and Crow, 1964; Turelli, 1984). For infinitely large populations and traits controlled by many independent loci with infinitesimal small effect, changes in trait optimum are not expected to affect the phenotypic variance (Lande, 1976). A much more complex picture is expected when the effect sizes are not equal, the population size is finite, or the traits have a simpler genetic basis (Barton and Keightley, 2002; Barton and Turelli, 1987; Franssen et al., 2017; Hayward and Sella, 2019; Jain and Stephan, 2015; Keightley and Hill, 1989). For instance, for a trait with oligogenic architectures, the genetic variance could drop dramatically during

53 adaptation, while with polygenic architectures, only minor effects on the variance are
54 expected (Barton et al., 2017; Franssen et al., 2017; Jain and Stephan , 2015). These studies
55 suggest that a time-resolved analysis of phenotypic variance has the potential to shed light
56 onto the complexity of the underlying adaptive architecture.
57
58 Despite its potential importance for the understanding of adaption, we are faced with the
59 situation that very few empirical data are available for the evolution of phenotypic variance.
60 The use of natural populations to study changes in phenotypes and even more so phenotypic
61 variances is limited, as the environmental heterogeneity cannot be controlled. A
62 complementary approach to study the evolution of phenotypic variance is experimental
63 evolution (Kawecki et al., 2012). With replicated populations starting from the same founders
64 and evolving under tightly controlled environmental conditions, experimental evolution
65 provides an enormous potential to study the evolution of phenotypic variance.
66
67 Most experimental evolution studies in sexual populations focused on the evolution of
68 phenotypic means, rather than variance (Burke et al., 2010; Chippindale et al., 1996; Jakšić et
69 al., 2020; Mallard et al., 2018). A notable exception is a study which applied fluctuating,
70 stabilizing and disruptive selection to a small number of traits (wing shape) and observed a
71 change of the phenotypic variance (Pélabon et al., 2010). Instead of looking at a preselected
72 subset of phenotypes which limits the generality, we will focus on gene expression, a set of
73 molecular phenotypes, which can be easily quantified since microarrays and, more recently,
74 RNA-Seq have become available. Importantly, the expression levels of genes exhibit the
75 same properties (e.g.: continuality and normality) as other complex quantitative traits
76 (Mackay et al., 2009). Thus, gene expression has also been widely employed to search for
77 putative adaptive traits of locally adapted populations (Romero et al., 2012; Signor and

Nuzhdin, 2018; Sork, 2017) or ancestral and evolved populations in the context of experimental evolution (Ferea et al., 1999; Huang and Agrawal, 2016; Lenski et al., 1994; Mallard et al., 2018).

In this study, we used forward simulations that match not only essential design features of typical experimental evolution studies, but also incorporate realistic parameters of the genetic architecture. We recapitulate the classic expectations that even a moderately polygenic architecture is associated with a high stability of the phenotypic variance of a selected trait across different phases of adaptation. Applying this insight to a real dataset (Barghi et al., 2019; Hsu et al., 2020; Jakšić et al., 2020), we show that a considerable set of genes changed their mean expression, but their expression variance was indistinguishable from genes without changes in mean expression. We propose that this pattern reflects that adaptive gene expression evolution generally has a polygenic basis.

Results

The central idea of this study is that the complexity of an adaptive trait can be estimated from the trajectory of the phenotypic variance during adaptation: the phenotypic variance remains relative stable for a trait with polygenic (infinitesimal) architecture while it changes across generations for a trait with oligogenic architecture. Although this prediction has been illustrated in multiple theoretical and simulation studies (e.g.: Barton et al., 2017; Franssen et al., 2017; Jain and Stephan, 2015), as the first step of this study, we explored to what extent it can be generalized to a typical E&R setting considering a broader parameter space. Assuming additivity and a negative correlation between ancestral allele frequency and effect size (Otte et al., 2020) (Figure 1b), we simulated traits adapting to a mild/distant shift in trait optimum with weak/intermediate/strong stabilizing selection (Figure 1a and Figure 1 – Figure

103 supplement 1). With three different distributions of effect size (Figure 1c) we investigate how
104 the number of contributing loci affects the phenotypic variance.
105
106 We monitored the change in phenotypic variance over 200 generations, which was sufficient
107 to reach the trait optimum for most parameter combinations (Figure 2 – Figure supplement 4
108 and 5). We compared the change in variance relative to the start of the experiment in
109 populations with and without selection. First, we studied a mild (one standard deviation of
110 the ancestral phenotypic distribution) shift in trait optimum (Figure 2 and Figure 2 – Figure
111 supplement 2). As expected for a founder population derived from a substantially larger
112 natural population, we find that even under neutrality the phenotypic variance does not
113 remain constant, but gradually decreases during 200 generations of experimental evolution
114 (Figure 2). This loss of variance is best explained by the fixation of variants segregating in
115 the founder population and we did not simulate new mutations, as they do not contribute to
116 adaptation in such short time scales (Burke et al., 2010). Our simulations show that even
117 experimental evolution studies with moderate population sizes and linkage very nicely
118 recapitulate the patterns from previous studies (Barton et al., 2017; Franssen et al., 2017; Jain
119 and Stephan , 2015). A pronounced drop in phenotypic variance is observed while a trait is
120 approaching a shifted optimum with few contributing loci (Figure 2). When more loci are
121 contributing to the selected phenotype, the difference to neutrality becomes very small
122 (Figure 2 and Figure 2 - Figure supplement 2). In addition to the number of contributing loci,
123 also the heterogeneity in effect size among loci and the shape of the fitness function have a
124 major impact. The larger the difference in effect size is, the more pronounced was the
125 influence of the number of contributing loci (Figure 2). The opposite effect was seen for the
126 width of the fitness function – a larger variance decreased the influence of the number of
127 contributing loci (Figure 2). Importantly, these patterns were not affected by the duration of

128 the experiment-qualitatively identical patterns were seen at different time points until
129 generation 200.
130
131 For a more distant trait optimum (three standard deviations of the ancestral phenotypic
132 distribution), we noticed some interesting dynamics that were not apparent for a closer trait
133 optimum (Figure 2 – Figure supplement 1 and 3). The most striking one was the temporal
134 heterogeneity of the phenotypic variance for few contributing loci of unequal effects. During
135 the early stage of adaptation, the variance increased and dropped later below the variance in
136 the founder population. With an increasing number of contributing loci, this pattern
137 disappeared and closely matched the neutral case (Figure 2 – Figure supplement 1).
138 Modifying dominance did not change the overall patterns-with a large number of contributing
139 loci the variance fitted the neutral pattern best (Figure 2 – Figure supplement 6). Overall, our
140 simulations indicate that only for a small set of parameters, the variance will increase during
141 the early stage of adaptation - in particular scenarios based on a few contributing loci with
142 very different effect sizes. The large influence of key parameters of the adaptive architecture,
143 in particular the number of contributing loci and their effect sizes on the temporal phenotypic
144 variance dynamics, suggest that it should be possible to exploit this for a test of polygenic
145 adaptation, which is independent from genomic data.
146
147 As a proof of principle, we studied the evolution of gene expression variance in replicated
148 populations evolving in a new hot temperature regime (Barghi et al., 2019; Hsu et al., 2020;
149 Jakšić et al., 2020). The evolved populations were derived from the same ancestral
150 population, but evolved independently for 100 generations in a novel temperature regime
151 with daily temperature fluctuations between 18 and 28°C (Figure 3a). Rather than relying on
152 pooled samples that allow only to estimate means, we quantified gene expression of

153 individuals from reconstituted ancestral populations and two evolved populations in a
154 common garden setup. Similar to previous studies (Ayroles et al., 2009; Huang et al., 2015),
155 we estimated a heritability of around 60% across the transcriptome among individual flies
156 from different families, which demonstrates the robustness of our experimental setup to
157 estimate expression variances (Method – Figure supplement 1; See Materials and methods).
158 Principle Component Analysis (PCA) indicated that 11.9% of the variation in gene
159 expression can be explained by the first PC which separates evolved and ancestral
160 populations, reflecting clear adaptive gene expression changes in response to the novel, hot
161 temperature regime (Figure 3b). The means and variances of the expression of each gene
162 were estimated and compared between the reconstituted ancestral populations and the two
163 evolved populations (Method – Figure supplement 2 and 3; See Materials and methods). Due
164 to the usage of different lot numbers for the RNA-Seq library preparation (Supplementary file
165 1), we only contrasted ancestral and evolved samples generated with the same lot number
166 (See Materials and methods) to avoid any unnecessary confounding results.
167
168 The comparison of ancestral and evolved populations identified 2,812 genes in the first
169 replicate and 2,704 genes in the second replicate which significantly changed mean
170 expression in the evolved flies (FDR<0.05, Supplementary file 2). With about 20% of the
171 genes changing mean expression, it is apparent that both populations evolved during 100
172 generations of exposure to a novel environment. 93.8% of the genes with a significant mean
173 expression change in both populations changed in the same direction, more than expected by
174 chance (Figure 4a, $\chi^2 = 896.34$, p-value $< 2.2 \times 10^{-16}$). This concordance suggests that
175 most of the altered expression means are mainly driven by selection, rather than by drift. We
176 quantified the expression change by relating the change in gene expression to the standard
177 deviation in the ancestral population. The differentially expressed genes in both replicates

178 showed a broad distribution of expression change, but the mean expression changed by one
179 standard deviation (Figure 4b). Assuming that all expression phenotypes reached trait
180 optimum, this reflects on average a moderate shift in trait optimum.
181
182 Consistent with computer simulations, the analysis of gene expression variance showed a
183 slight decrease in evolved populations relative to the ancestral ones (median F-value = 0.84 in
184 both replicates). Only a small subset of gene (n=125 and 97 in each replicate) experienced a
185 substantial reduction in variance. Because this reduction in variance is probably driven by a
186 different evolutionary force, we discuss them elsewhere (Lai et al., 2021). For the remaining
187 genes, we related the changes in gene expression variance of selected genes to non-selected
188 ones, we tested whether the variance changes in expression differ between the genes with
189 significant mean expression changes and those without. Hence, we assume that genes with
190 significant mean expression changes are under selection and the rest of the transcriptome has
191 no large effect on the fitness (neutral). Remarkably, the changes in variance of putative
192 adaptive genes with significant mean expression changes are indistinguishable from the genes
193 that do not change their mean expression (Figure 4c). This suggests that the selection on
194 mean expression is independent from the change in variance during adaptation in the focal
195 populations. While variance estimates from two time points do not provide sufficient power
196 to estimate the number of contributing loci in absence of more information about the adaptive
197 architecture, our computer simulations (Figure 2), suggest in line with previous theoretical
198 work (Barton et al., 2017; Jain and Stephan, 2015) that the observed stability in variance
199 evolution reflects a polygenic architecture underlying the adaptive gene expression evolution.
200
201 Since we only explored two time points rather than a full time series, it may be possible that
202 an oligogenic basis could also result in a similar phenotypic variance change as a polygenic

203 architecture (Figure 2 – Figure supplement 1 and 3). This can be seen in an intuitive case
204 when a single/few major effect allele(s) starts at a low frequency and becomes fixed (Yoo et
205 al., 1980). Because an oligogenic basis results in a highly parallel genomic selection response
206 (Figure 2 – Figure supplement 7), it is possible to distinguish polygenic and oligogenic
207 architectures with phenotypes from two time points only, when genomic data are available.
208 Because the genomic signature in the same experiment uncovered a highly heterogeneous
209 selection response (Barghi et al., 2019), we can exclude the unlikely explanation of an
210 oligogenic architecture resulting in a similar expression variance as non-selected genes.
211 Rather, we conclude that the adaptive response in gene expression is best explained by a
212 highly polygenic architecture.

Discussion

Population genetics has a long tradition to characterize adaptation based on the genomic signature of selected loci (Nielsen, 2005). Nevertheless, for selected phenotypes with a polygenic architecture, the contribution of individual loci to phenotypic change may be too subtle to be detected with classic population genetic methods (Pritchard et al., 2010). Even with an oligogenic basis, the identification of the selection targets with classic population genetic tests can be challenging.

Here, we used a conceptually different approach, which does not build on the genomic signature, to infer the adaptive architecture. Reasoning that experimental evolution is probably the best approach to obtain phenotypic time series, we performed computer simulations specifically tailored to typical experimental evolution studies with *Drosophila*. We showed that the temporal dynamics of the phenotypic variance is strongly affected by the number of contributing loci and other parameters of the adaptive architecture, such as the distribution of effect size and the underlying fitness function. Similar to the classic Castle-Wright estimator (Castle, 1921) that estimates the number of loci contributing to a quantitative trait from the phenotypic variance of the F2, we propose that the temporal heterogeneity of the phenotypic variance can be used to estimate the number of loci contributing to the adaptive response of a phenotype as well as other parameters of the adaptive architecture. Hence, unlike other approaches to characterize polygenic adaptation, the proposed estimator does not require genetic data when phenotypic time series data are available.

Because gene expression changes are constituting a major component of adaptation to a novel environment (Romero et al., 2012; Signor and Nuzhdin, 2018; Sork, 2017), it provides an

238 excellent model to evaluate a variance-based test for polygenicity. Gene expression is
239 modified by many trans-acting factors and some cis-regulatory variation. Adaptive gene
240 expression changes can be either driven by polymorphism in *cis*-regulatory elements or by
241 *trans*-acting variants. While interspecific differences in gene expression are predominantly
242 caused by *cis*-regulation (Wittkopp et al., 2004), intraspecific variation is mostly driven by
243 trans regulatory changes (Suvorov et al., 2013; Wittkopp et al., 2008). Adaptive gene
244 expression changes which are well-characterized on the molecular level typically have a
245 *cis*-regulatory basis that is not only frequently associated with the insertion of a transposable
246 element (e.g.: (Daborn et al., 2002)) but also sometimes with multiple regulatory variants
247 (Endler et al., 2018). Two lines of evidence suggest that cis-regulatory variation cannot be the
248 driver of adaptive gene expression changes observed in this study. First, the mutational target
249 size is too small to harbor a sufficiently large number of alleles segregating in the founder
250 population. Second, too few recombination events occur during the experiment to uncouple
251 regulatory variants located on a given haplotype such that they could generate a signal of
252 polygenic adaptation. More likely, the polygenic adaptive architecture of gene expression
253 change reflects the joint effects of many *trans*-acting variants.
254
255 Because we could only analyze phenotypic data from two time points, the founder population
256 and replicate populations evolved for 100 generations, we were not able to obtain a more
257 quantitative estimate of the number of contributing loci, in particular as other parameters of
258 the adaptive architecture are not known and need to be co-estimated. For the distinction
259 between an oligogenic and polygenic basis, we additionally relied on the heterogeneity of
260 genomic selection signatures among replicates, because for some parameter combinations the
261 oligogenic response can also result in a similar phenotypic variance as a polygenic one, but
262 with a much higher parallel response of genomic markers (Figure 2 – Figure supplement 7).

263 Hence, not only more time points describing the phenotypic trajectory, but also some
264 genomic data could contribute to infer the adaptive architecture in experimental evolution
265 studies. The extension of this approach to natural populations faces several challenges. First,
266 phenotypic time series over evolutionary relevant time scales are rare (but see (Clutton-Brock
267 and Pemberton, 2004)) and second, the distinction of environmental heterogeneity from
268 genetic changes is considerably more challenging than under controlled laboratory
269 conditions.
270
271

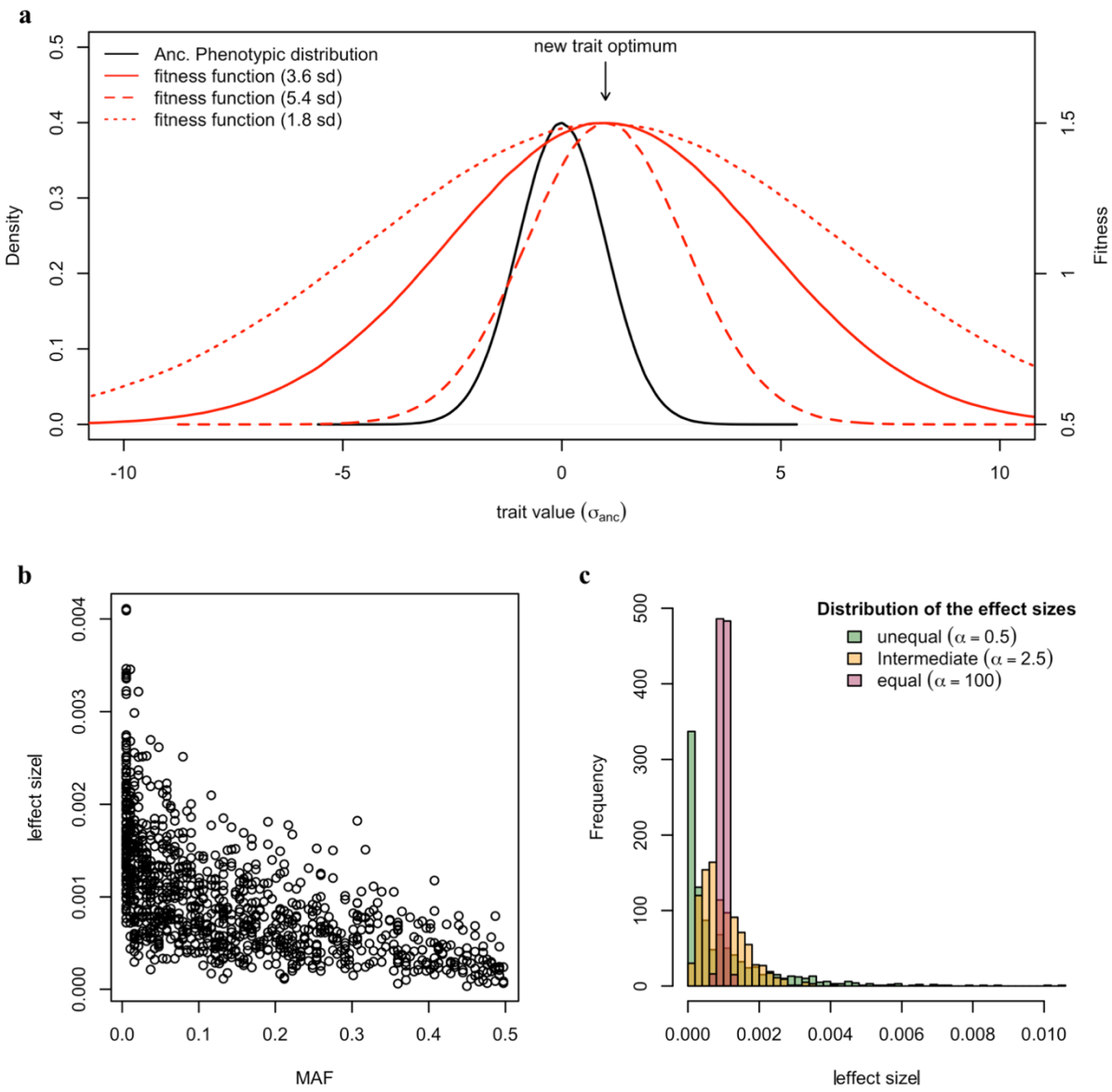


Figure 1. Simulating polygenic adaptation to a shift in trait optimum with different parameter combinations. a. For the computer simulations we consider a quantitative trait (in black) experiencing a sudden shift in trait optimum under stabilizing selection. The underlying fitness functions are illustrated in red. The new trait optimum is shifted from the ancestral trait mean by one/three standard deviation of the ancestral trait distribution. The strength of stabilizing selection is modified by changing the variance of the fitness function: 1.8, 3.6 and 5.4 standard deviations of the ancestral trait distribution. **b.** The negative correlation between the allele frequencies and the effect sizes (r = -0.7, estimated in Barghi *et*

281 *al*., 2019). We consider such negative correlation when assigning the effect sizes to variants
282 underlying a simulated trait. **c.** The distribution of effect sizes of the contributing loci is
283 determined by the shape parameter (α) of gamma sampling process (α = 0.5, 2.5 and 100).
284

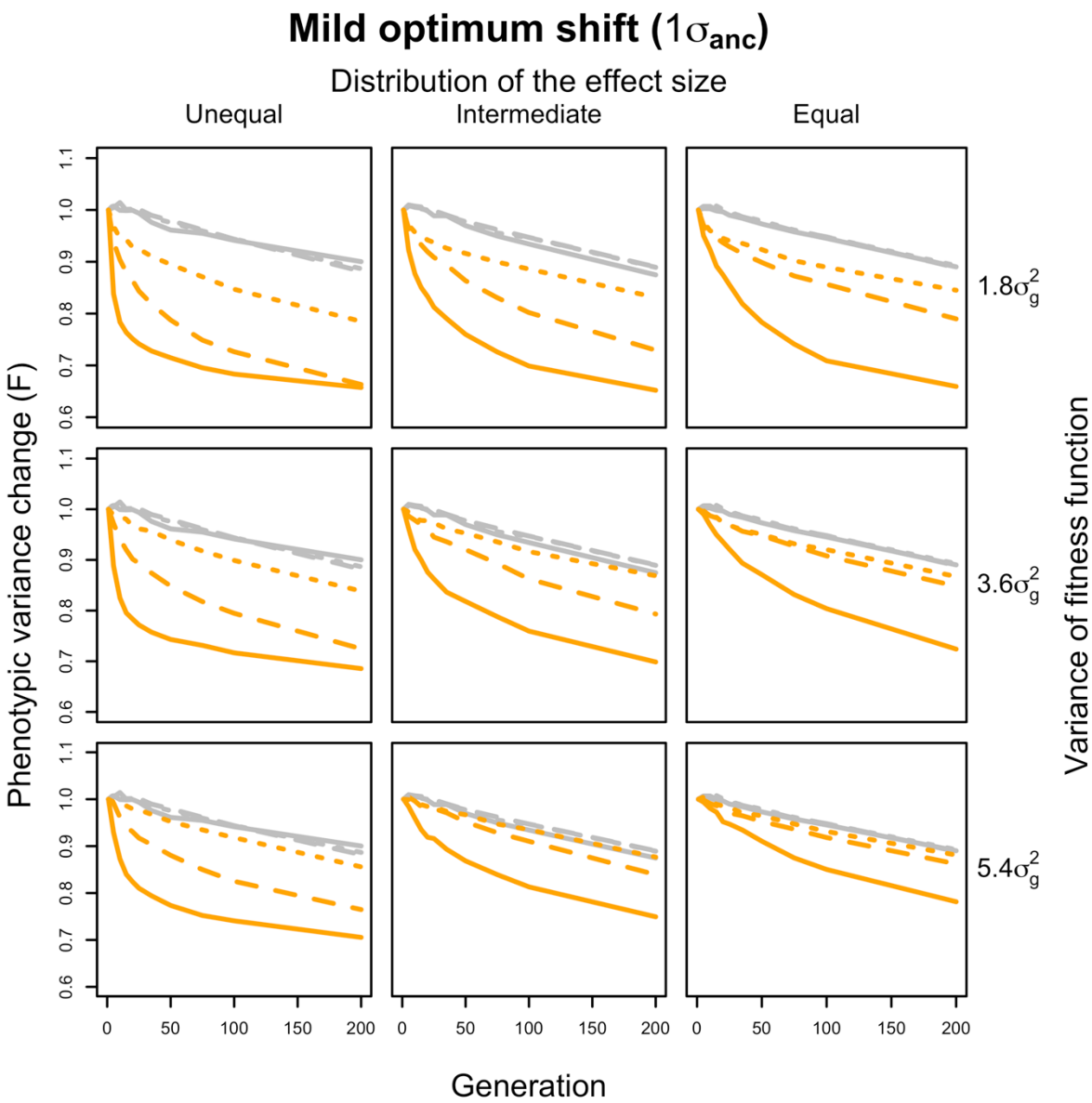


Figure 2. The trajectory of changes in phenotypic variance during adaptation to a mild optimum shift. The changes in phenotypic variance within 200 generations adapting to a moderate optimum shift (orange) are compared to the changes under neutrality (grey) on y axis. The change in variance (F) is calculated as the ratio of phenotypic variance between each evolved time point (generation x) and the ancestral state (σ_x^2/σ_0^2). The simulations cover traits controlled by varying numbers of loci underlying the adaptation with three different distributions of effect sizes (columns) under different strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations have been performed. Only traits with the most

(dotted lines, 1000 loci), intermediate (dash lines, 50 loci) and the least (solid lines, 5 loci) polygenic architectures are shown. In all scenarios, the variance of the trait decreases drastically when the adaptation is controlled by a small number of loci (orange solid lines; 5 loci). While, for traits with extremely polygenic basis, the phenotypic variance stays stable over time (orange dotted lines).

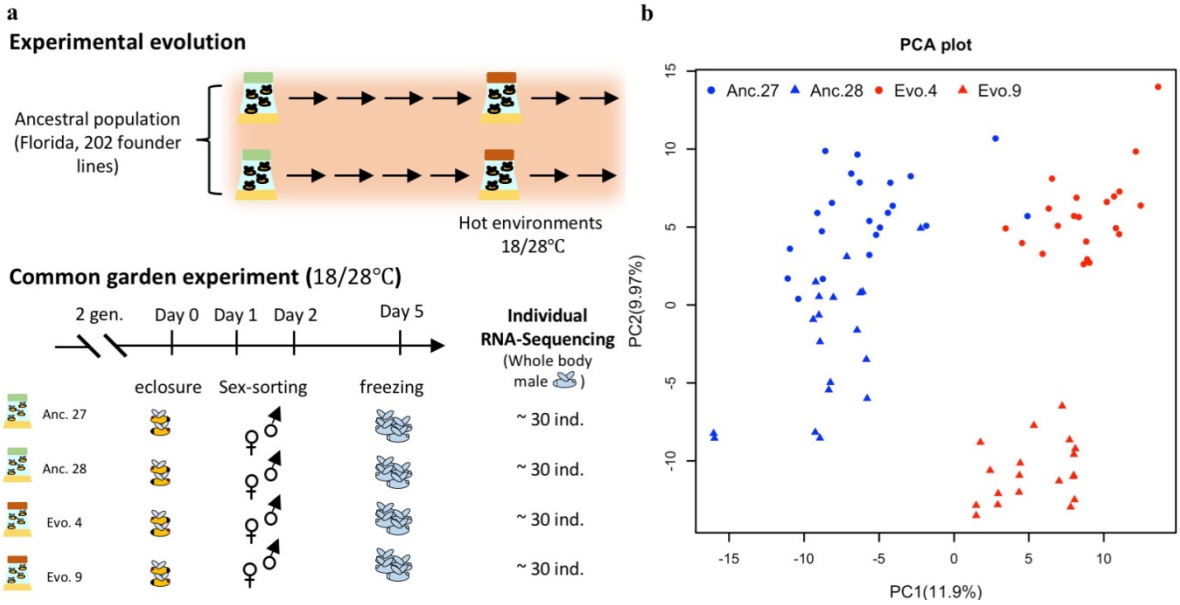
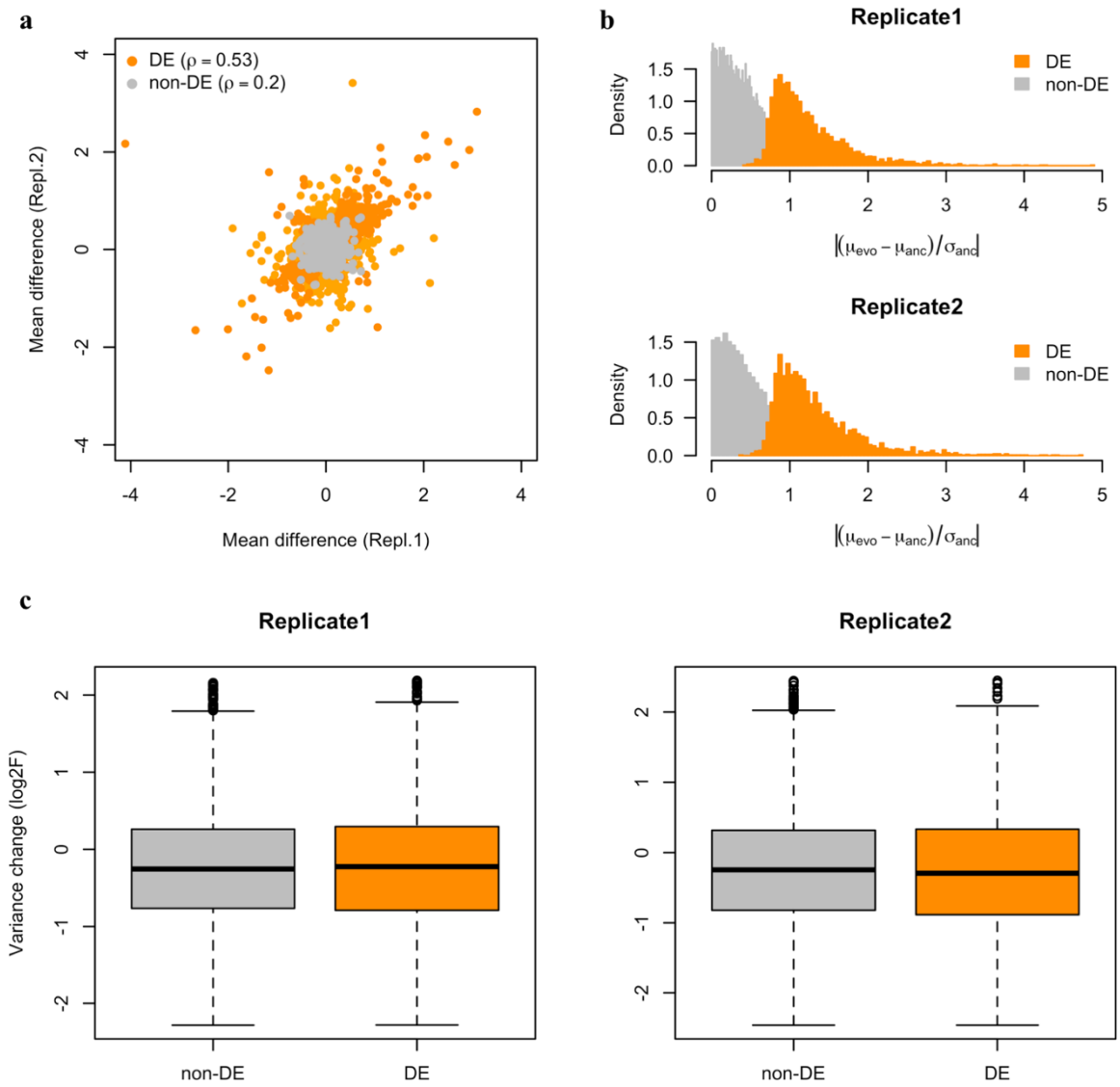


Figure 3. Schematic overview of the experimental procedures (a) and the divergence in gene expression during experimental evolution (b). **a.** Experimental evolution: starting from one common founder population, two replicate populations evolved for 100 generations in a hot laboratory environment fluctuating between 18 and 28°C. Common garden Experiment: after 100 generations, the two evolved replicate populations were maintained together with the reconstituted ancestral population for two generations in a hot laboratory environment fluctuating between 18 and 28°C. After this common garden procedure, about 30 males from each population were subjected to RNA-Seq. **b.** Principle Component Analysis (PCA) of the transcriptomic profiles of individuals from the ancestral population (blue) and the hot-evolved population (red). Circles indicate individuals of the first replicate (Anc. No. 27 and Evo. No. 4). Triangles represent individuals of the second replicate (Anc. No. 28 and Evo. No. 9). The two replicates were made with two different batches of library cards for RNA-Seq library preparation.



315

316 **Figure 4. Evolution of phenotypic mean and variance over 100 generations of**
317 **adaptation in empirical data. a.** The evolution of gene expression mean during adaptation
318 in the two replicates. For the genes with significant changes (DE, in orange), the changes are
319 correlated between replicates (Spearman's rho = 0.53). For the genes without significant
320 changes (non-DE, in grey), the correlation between replicates is much lower (Spearman's rho
321 = 0.2). **b.** The evolution of gene expression mean scaled by the ancestral variation. For the
322 DE genes (in orange), the median change falls around 1 standard deviation of the ancestral
323 expression value, suggesting mild shift in trait optimum in the novel environment. For the

324 non-DE genes (in grey), the changes in expression are mostly negligible. The same pattern is
325 seen in the second replicate. **c.** The change in expression variance during adaptation for DE
326 and non-DE genes. In both replicates, the distribution of variance changes is
327 indistinguishable between DE genes (orange) and non-DE genes (grey) (Wilcoxon's rank
328 sum test, $p > 0.1$ for both replicates).
329

330 **Materials and Methods**

331 Computer simulations

332 We performed forward simulations with MimicrEE2 (Vlachos and Kofler, 2018) using the
333 qff mode to illustrate the influence of the genetic architecture on the evolution of phenotypic
334 variance during the adaptation to a new trait optimum. With 189 founder haplotypes (Barghi
335 et al., 2019), we simulated quantitative traits under the control of different numbers of loci (5,
336 25, 50, 100, 200 and 1000) with an effective population size of 300. For each trait, we
337 assume an additive model and a negative correlation (r = -0.7, estimated in Barghi et al.,
338 2019) between the effect size and starting frequency (Figure 1b). The effect sizes of each
339 locus can disperse in different levels which depend on the shape parameter of gamma
340 sampling process (shape = 0.5, 2.5 and 100, Figure 1c). We used correlate() function
341 implemented in "fabricatr" R package (Blair et al., 2019) to generate the effect sizes. The
342 sum of effect sizes of each trait was normalized to 1. We assumed a heritability $h^2 = 0.6$
343 (from a family-based estimation in this study). To simulate stabilizing selection with trait
344 optimum shift, we provided the Guassian fitness functions with mean of $\overline{X_{anc.}} + a\sqrt{V_{anc.}}$
345 and standard deviation of $b\sqrt{V_{anc.}}$, where $\overline{X_{anc.}}$ is the ancestral phenotypic mean and $V_{anc.}$
346 is the ancestral genetic variance (Figure 1a). Parameter "a" determines the distance of
347 optimum shift, which is set to one (similar to the empirical case, Figure 4b) or three (Adopted
348 from Sella et al., 2019). Parameter "b" indicates how strong the phenotypic constrain would
349 be when the trait optimum is reached. The value 3.6 is adopted from Sella et al., 2019. In this
350 study, we increase and decrease it by 50% to explore its impact (1.8 or 5.4). For the neutrality
351 case, we assumed uniform fitness for all individuals. For each trait under each scenario, the
352 phenotypic variance was estimated at different generations and compared to the ancestral
353 phenotypic variance at generation 1 to illustrate the dynamic of phenotypic variance during
354 the evolution. We note that we do not assume that the ancestral population has reached an

355 equilibrium, because the ancestral population in this study was phenotyped in the new
356 environment.
357
358 Experimental evolution
359 The setup of populations and evolution experiment have been described by (Barghi et al.,
360 2019). Briefly, ten outbred populations seeded from 202 isofemale lines were exposed to a
361 laboratory experiment at 28/18 °C with 12hr light/12hr dark photoperiod for more than 100
362 generations. Each replicate consisted of 1000 to 1250 adults at each generation.
363
364 Common garden experiment
365 The collection of samples from the evolution experiment for RNA-Seq was preceded by two
366 generations of common garden (CGE). The common garden experiment was performed at
367 generation 103 of the evolution in the hot environment and this CGE has been described in
368 (Barghi et al., 2019; Hsu et al., 2020, 2019; Jakšić et al., 2020). In brief, an ancestral
369 population was reconstituted by pooling five mated females from 184 founder isofemale lines
370 (Nouhaud et al., 2016). No significant allele frequency differences are expected between the
371 reconstituted ancestral populations and the original ancestral populations initiating the
372 experiment (Nouhaud et al., 2016). Because we evaluated phenotypes on the population
373 level, deleterious mutations will have a very limited impact. The reason is that they occur
374 only in a single isofemale line, which represents a very small fraction of the total population.
375 Two replicates of the reconstituted ancestral population and two independently evolved
376 populations at generation 103 were reared for two generations with egg-density control (400
377 eggs/bottle) at the same temperature regime as in the evolution experiment. Freshly eclosed
378 flies were transferred onto new food for mating. Sexes were separated under CO_2 anesthesia
379 at day 3 after eclosure, left to recover from CO_2 for two days, and at the age of five days

380 whole-body mated flies of each sex were snap-frozen at 2pm in liquid nitrogen and stored at
381 -80°C until RNA extraction. In this study, more than 30 individual male flies from two
382 reconstituted ancestral populations (replicate no. 27 and no. 28) and two evolved populations
383 (replicate no. 4 and no. 9) were subjected to RNA-Seq.
384
385 RNA extraction and library preparation
386 Whole bodies of individual male flies were removed from the -80°C freezer and immediately
387 homogenized in Qiazol lysis reagent (Qiagen, Hilden, Germany). The homogenate was
388 treated with DNase I followed by addition of chloroform, centrifugation and mixture of the
389 upper phase with 70% ethanol as described for the Qiagen RNeasy Universal Plus Mini Kit.
390 The mixture was subsequently loaded onto a RNeasy MinElute Spin column as provided by
391 the RNeasy Plus Micro Kit (Qiagen, Hilden, Germany), and all washing steps were
392 performed according to the instructions for that kit. All resulting total RNA was used to
393 prepare stranded mRNA libraries on the Neoprep Library Prep System (Illumina, San Diego,
394 USA) following the manufacturer's protocol: Neoprep runs were performed using software
395 version 1.1.0.8 and protocol version 1.1.7.6 with default settings for 15 PCR cycles and an
396 insert size of 200bp. All libraries for individuals of ancestral replicate no. 27 and evolved
397 replicate no. 4 were prepared with library cards of lot no. 20180170; all libraries for
398 individuals of ancestral replicate no. 28 and evolved replicate no. 9 were prepared with
399 library cards of lot no. 20178099. 50bp single-end reads were sequenced on an Illumina
400 HiSeq 2500. All sequencing data will be available in European Nucleotide Archive (ENA)
401 under the accession number PRJEB37011 upon publication.
402
403 RNA-Seq data processing and quality control

404 All RNA-Seq reads were trimmed using ReadTools (Gómez-Sánchez and Schlötterer, 2018)
405 with quality score of 20 and aligned to *Drosophila simulans* reference genome (Palmieri et
406 al., 2015) using GSNAP (Wu et al., 2016) with parameter setting -k 15 -N 1 -m 0.08.
407 Exon-aligned reads were piped into Rsubread (Liao et al., 2019) to calculate read counts of
408 each gene, and raw read counts of each gene were normalized with the TMM method
409 implemented in edgeR (Robinson et al., 2010). Samples with severe 3'- bias were removed
410 based on visual inspection of the gene-body coverage plot (Jakšić and Schlötterer, 2016;
411 Wang et al., 2012).
412
413 Genetic variance in gene expression across F1 families
414 We evaluated how much of the expression variance can be explained by genetic variation by
415 performing RNA-Seq on individual flies, with 3-4 individuals each from three isofemale lines
416 maintained at the same density and culturing conditions.
417 Six out of 184 founder isofemale lines from the evolution experiment and were maintained
418 for one generation with controlled egg density (400 eggs/bottle) in the same environment as
419 the main experiment (12h 28°C with light followed by 12h 18 °C with dark conditions).
420 Using the offspring, we generated three crosses between two of the six lines each: FL 138 x
421 FL 137, FL 157 x FL 112, FL 123 x FL 127: we combined 50 virgin females from one of the
422 lines with 50 males from the other line, let them lay eggs under density control as above and
423 maintained and froze their F1 offspring in the same way as in the main CGE: sexes were
424 separated after mating at the age of three days and snap-frozen at the age of five days at 2pm.
425 From each cross, we used four F1 males to prepare individual RNA-Seq libraries as described
426 above.

427 Assuming no environmental heterogeneity, we decomposed the total variance of the
428 expression of each gene measured in these individuals into the genetic difference among
429 three different F1 families and random error. The data were analyzed as follows:
430 Natural log-transformation was applied to CPMs of all genes to improve data normality
431 (Rocke and Durbin, 2003). Principal component and principal variance component analysis
432 (Bushel, 2019) was performed to the whole transcriptome to decompose the variance
433 components. We found that around 60% of the gene expression variance can be explained by
434 the genetic difference among the three F1 families (method – Figure supplement 1a and b).
435 This implies that the within-population gene expression variance is largely contributed by the
436 genetic components. Because we only used offspring from single vials, we may have
437 overestimated the heritability if the environment in the vials differs. Nevertheless, since our
438 heritability estimates are very similar to previous ones (Ayroles et al., 2009), we consider our
439 estimates reliable.
440
441 In addition to general analysis across all genes, we also tested for the genetic variance of each
442 gene separately using analysis of variance (ANOVA):
443 $$y_{ij} = \mu + \tau_i + \varepsilon_{ij},$$
444 Where i=1, 2, 3, …, n (the ith genes); j =1, 2, 3, 4 (the jth individuals in each cross). y_{ij} is the
445 observed expression level of a gene in a given sample, μ is the overall mean; τ_i is the
446 effect of genetic background and ε_{ij} is the random noise. We calculated the proportion of
447 total variance explained by random error using the following equation:
448 $$variance\ explained\ by\ random\ error = \frac{sum\ squares\ of\ error\ (SSE)}{sum\ squares\ of\ total\ (SST)}$$

449 Genes were binned based on their average expression value (lnCPM) which ranged from -0.8
450 to 4.1, by bin size of 0.1. The average proportion of variance explained by random error of
451 each bin was calculated.
452 The expression variance of genes with less than 1 count per million bases (CPM) is
453 dominated by technical errors (method – Figure supplement 1c). Thus, genes with less than
454 1 count per million base (CPM) were excluded for subsequent analysis.
455
456 RNA-Seq data analysis
457 We observed some outlier individuals and suspected that the freezing process may have led
458 to detachment of body parts, such as eyes or heads, in these individuals. We compared gene
459 expression between such outliers and all other samples and performed tissue enrichment
460 analysis for genes with at least 2-fold lower expression in the outlier samples. Samples with
461 evidence of tissue detachment were excluded. After filtering, each population remained
462 approximately 20 individuals (Supplementary file 1). Only genes with at least 1 count per
463 million base (CPM) were included in the analyses to avoid extremely lowly expressed genes.
464 For all RNA-Seq data we only compared samples which were prepared with library cards
465 from the same lot number to avoid batch effects (Replicate 1: evolved replicate 4 vs.
466 reconstituted ancestral population replicate 27; Replicate 2: evolved replicate 9 ss.
467 reconstituted ancestral population replicate 28).
468 For differential expression analysis on mean expression, we used the generalized linear
469 modeling function implemented in edgeR (Robinson et al., 2010) to fit the expression to the
470 model ($Y = E + \varepsilon$) in which Y stands for gene expression, E is the effect of evolution and ε
471 is the random error. The likelihood ratio test was performed to test the effect of evolution.
472 P-value adjustment was performed using the Benjamini-Hochberg false discovery rate (FDR)
473 correction.

474 For the analysis of expression variance evolution, we applied natural log transformation
475 (Rocke and Durbin, 2003) to eliminate the strong mean-variance dependency in RNA-Seq
476 data due to the nature of the negative binomial distribution (method – Figure supplement 2).
477 The variance of the expression of each gene (lnCPM) was estimated in each population. With
478 the moderate sample size, we needed to estimate the uncertainty of variance estimates.
479 Jackknifing was applied to measure the uncertainty of estimator (Fukunaga and Hummels,
480 1989). The procedure was conducted independently on four replicates and we calculated the
481 95% confidence interval of the estimated variance (method – Figure supplement 3). The
482 change of gene expression variance was determined by the F statistics calculated as the ratio
483 between the variance within the ancestral population and the variance within the evolved
484 population of each gene. To test whether selection alter the expression variance, a
485 comparison was made between the F statistics of genes with significant changes in mean
486 expression and the ones without.
487

Acknowledgments

Special thanks to David Houle, who provided fantastic support during the collection and establishment of the isofemale lines in Florida. We thank all member of the Institut für Populationsgenetik for discussion. We are grateful to Reinhard Bürger and David Houle for helpful comments on earlier versions of the manuscript. Neda Barghi, François Mallard and Kathrin Otte contributed to the common garden experiment. Illumina sequencing was performed at the VBCF NGS Unit (www.vbcf.ac.at). This work was support by the Austrian Science Funds (FWF, W1225) and the European Research Council (ERC, ArchAdapt).

Author contribution

W.Y.L and C.S. conceived the study. V.N. prepared all RNA-Seq and supervised the maintenance of the evolution experiment. A.M.J supervised the common garden experiment. W.Y.L performed the simulation and data analysis. W.Y.L. and C.S. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Correspondence and requests for materials should be addressed to C.S.

507

508 **References**

509 Ayroles JF, Carbone MA, Stone EA, Jordan KW, Lyman RF, Magwire MM, Rollmann SM,
510 Duncan LH, Lawrence F, Anholt RRH, Mackay TFC. 2009. Systems genetics of
511 complex traits in *Drosophila melanogaster*. *Nat Genet* **41**:299–307. doi:10.1038/ng.332

512 Barghi N, Hermisson J, Schlötterer C. 2020. Polygenic adaptation: a unifying framework to
513 understand positive selection. *Nat Rev Genet* **21**:769–781.
514 doi:10.1038/s41576-020-0250-z

515 Barghi N, Tobler R, Nolte V, Jakšić AM, Mallard F, Otte KA, Dolezal M, Taus T, Kofler R,
516 Schlötterer C. 2019. Genetic redundancy fuels polygenic adaptation in *Drosophila*.
517 *PLOS Biol* **17**:e3000128. doi:10.1371/journal.pbio.3000128

518 Barton NH, Etheridge AM, Véber A. 2017. The infinitesimal model: Definition, derivation,
519 and implications. *Theor Popul Biol* **118**:50–73. doi:10.1016/J.TPB.2017.06.001

520 Barton NH, Keightley PD. 2002. Understanding quantitative genetic variation. *Nat Rev Genet*
521 **3**:11–21. doi:10.1038/nrg700

522 Barton NH, Turelli M. 1987. Adaptive landscapes, genetic distance and the evolution of
523 quantitative characters. *Genet Res* **49**:157–173. doi:10.1017/S0016672300026951

524 Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: From polygenic
525 to omnigenic. *Cell* **169**:1177–1186. doi:10.1016/J.CELL.2017.05.038

526 Bulmer MG. 1972. The genetic variability of polygenic characters under optimizing
527 selection, mutation and drift. *Genet Res* **19**:17–25. doi:10.1017/S0016672300014221

528 Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. 2010.
529 Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature*
530 **467**:587–590. doi:10.1038/nature09352

531 Bushel P. 2019. pvca: Principal variance component analysis (PVCA). *R Packag version*

532 *1240*.

533 Castle WE. 1921. An improved method of estimating the number of genetic factors
534 concerned in the cases of blending inheritance. *Science* **54**:223.
535 doi:10.1126/science.54.1393.223

536 Chevalet C. 1994. An approximate theory of selection assuming a finite number of
537 quantitative trait loci. *Genet Sel Evol.* **26**:379. doi:10.1186/1297-9686-26-5-379

538 Chippindale AK, Chu TJF, Rose MR. 1996. Complex trade-offs and the evolution of
539 starvation resistance in *Drosophila melanogaster*. *Evolution (N Y)* **50**:753–766.
540 doi:10.1111/j.1558-5646.1996.tb03885.x

541 Clutton-Brock TH, Pemberton JM (Josephine M). 2004. Soay sheep : population dynamics
542 and selection on St. Kilda. Cambridge University Press.

543 Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D,
544 Batterham P, Feyereisen R, Wilson TG, ffrench-Constant RH. 2002. A single p450
545 allele associated with insecticide resistance in *Drosophila*. *Science* **297**:2253–6.
546 doi:10.1126/science.1074170

547 Endler L, Gibert J-M, Nolte V, Schlötterer C. 2018. Pleiotropic effects of regulatory variation
548 in *tan* result in correlation of two pigmentation traits in *Drosophila melanogaster*. *Mol*
549 *Ecol* **27**:3207–3218. doi:10.1111/mec.14781

550 Falconer DS, Mackay TFC. 1963. Introduction to quantitative genetics, Poultry Science.
551 CreateSpace Independent Publishing Platform. doi:10.3382/ps.0420547

552 Ferea TL, Botstein D, Brown PO, Rosenzweig RF. 1999. Systematic changes in gene
553 expression patterns following adaptive evolution in yeast. *Proc Natl Acad Sci U S A*
554 **96**:9721–6. doi:10.1073/pnas.96.17.9721

555 Franssen SU, Kofler R, Schlötterer C. 2017. Uncovering the genetic signature of quantitative
556 trait evolution with replicated time series data. *Heredity (Edinb)* **118**:42–51.

557 doi:10.1038/hdy.2016.98

558 Fukunaga K, Hummels DM. 1989. Leave-one-out procedures for nonparametric error
559 estimates. *IEEE Trans Pattern Anal Mach Intell* **11**:421–423. doi:10.1109/34.19039

560 Gómez-Sánchez D, Schlötterer C. 2018. *ReadTools*: A universal toolkit for handling
561 sequence data from different sequencing platforms. *Mol Ecol Resour* **18**:676–680.
562 doi:10.1111/1755-0998.12741

563 Hayward LK, Sella G. 2019. Polygenic adaptation after a sudden change in environment.
564 *bioRxiv* 792952. doi:10.1101/792952

565 Hsu S-K, Jakšić AM, Nolte V, Barghi N, Mallard F, Otte KA, Schlötterer C, Hsu S-K, Jakšić
566 AM, Nolte V, Barghi N, Mallard F, Otte KA, Schlötterer C. 2019. A 24 h age difference
567 causes twice as much gene expression divergence as 100 generations of adaptation to a
568 novel environment. *Genes (Basel)* **10**:89. doi:10.3390/genes10020089

569 Hsu S-K, Jakšić AM, Nolte V, Lirakis M, Kofler R, Barghi N, Versace E, Schlötterer C.
570 2020. Rapid sex-specific adaptation to high temperature in *Drosophila*. *Elife* **9**.
571 doi:10.7554/eLife.53237

572 Huang W, Carbone MA, Magwire MM, Peiffer JA, Lyman RF, Stone EA, Anholt RRH,
573 Mackay TFC. 2015. Genetic basis of transcriptome diversity in *Drosophila*
574 *melanogaster*. *Proc Natl Acad Sci U S A* **112**:E6010–E6019.
575 doi:10.1073/pnas.1519159112

576 Huang Y, Agrawal AF. 2016. Experimental evolution of gene expression and plasticity in
577 alternative selective regimes. *PLOS Genet* **12**:e1006336.
578 doi:10.1371/journal.pgen.1006336

579 Jain K, Stephan W. 2015. Response of polygenic traits under stabilizing selection and
580 mutation when loci have unequal effects. *G3 (Bethesda)* **5**:1065–74.
581 doi:10.1534/g3.115.017970

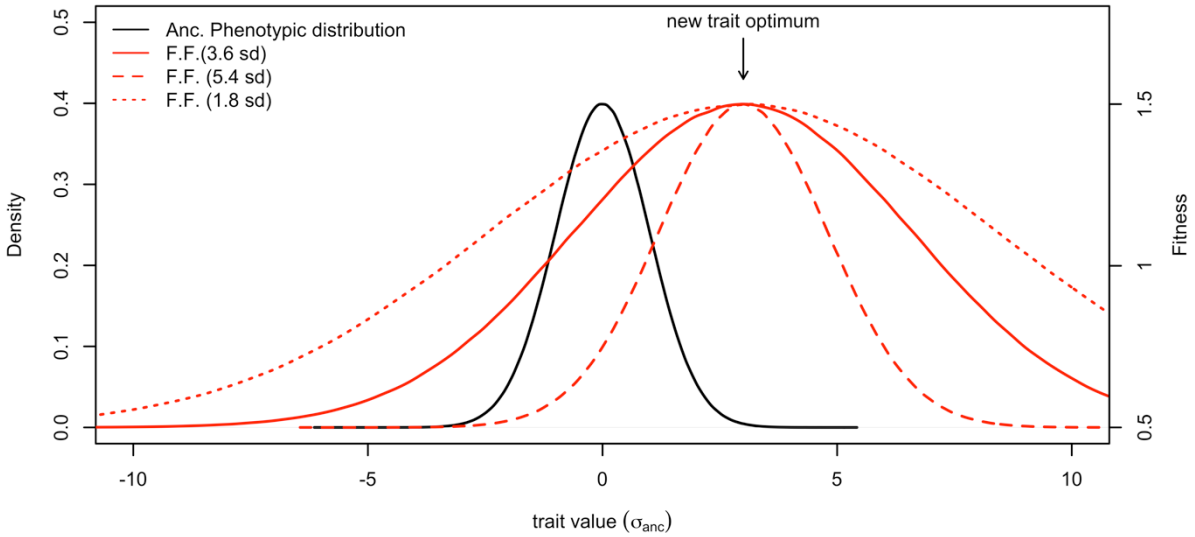
582 Jakšić AM, Karner J, Nolte V, Hsu S-K, Barghi N, Mallard F, Otte KA, Svečnjak L, Senti
583 K-A, Schlötterer C. 2020. Neuronal function and dopamine signaling evolve at high
584 temperature in *Drosophila*. *Mol Biol Evol* **37**:2630–2640. doi:10.1093/molbev/msaa116
585 Jakšić AM, Schlötterer C. 2016. The interplay of temperature and genotype on patterns of
586 alternative splicing in *Drosophila melanogaster*. *Genetics* **204**:315–325.
587 doi:10.1534/GENETICS.116.192310
588 Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012. Experimental
589 evolution. *Trends Ecol Evol* **27**:547–560. doi:10.1016/J.TREE.2012.06.001
590 Keightley PD, Hill WG. 1989. Quantitative genetic variability maintained by
591 mutation-stabilizing selection balance: sampling variation and response to subsequent
592 directional selection. *Genet Res* **54**:45–58. doi:10.1017/S0016672300028366
593 Kimura M, Crow JF. 1964. The number of alleles that can be maintained in a finite
594 population. *Genetics* **49**:725–738.
595 Lai W.-Y. & Schlötterer, C. 2021. Evolution of gene expression variance during adaptation to
596 high temperature in *Drosophila*. *BioRxiv*.
597 Lande R. 1976. Natural selection and random genetic drift in phenotypic evoltion. *Evolution*
598 *(N Y)* **30**:314–334. doi:10.1111/j.1558-5646.1976.tb00911.x
599 Lenski RE, Travisano M, Larison B, Moritz C. 1994. Dynamics of adaptation and
600 diversification: a 10,000-generation experiment with bacterial populations. *Proc Natl*
601 *Acad Sci* **91**:6808–6814. doi:10.1073/pnas.91.15.6808
602 Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and
603 better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*
604 **47**:e47–e47. doi:10.1093/nar/gkz114
605 Liu X, Li YI, Pritchard JK. 2019. Trans effects on gene expression can drive omnigenic
606 inheritance. *Cell* **177**:1022-1034.e6. doi:10.1016/J.CELL.2019.04.014

607 Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and
608 prospects. *Nat Rev Genet* **10**:565–577. doi:10.1038/nrg2612
609 Mallard F, Nolte V, Tobler R, Kapun M, Schlötterer C. 2018. A simple genetic basis of
610 adaptation to a novel thermal environment results in complex metabolic rewiring in
611 Drosophila. *Genome Biol* **19**:119. doi:10.1186/s13059-018-1503-4
612 Nielsen R. 2005. Molecular signatures of natural selection. *Annu Rev Genet* **39**:197–218.
613 doi:10.1146/annurev.genet.39.073003.112420
614 Nouhaud P, Tobler R, Nolte V, Schlötterer C. 2016. Ancestral population reconstitution from
615 isofemale lines as a tool for experimental evolution. *Ecol Evol* **6**:7169–7175.
616 doi:10.1002/ece3.2402
617 Otte KA, Nolte V, Mallard F, Schlötterer C. 2020. The adaptive architecture is shaped by
618 population ancestry and not by selection regime. *bioRxiv* 2020.06.25.170878.
619 doi:10.1101/2020.06.25.170878
620 Palmieri N, Nolte V, Chen J, Schlötterer C. 2015. Genome assembly and annotation of a
621 *Drosophila simulans* strain from Madagascar. *Mol Ecol Resour* **15**:372–81.
622 doi:10.1111/1755-0998.12297
623 Pélabon C, Hansen TF, Carter AJR, Houle D. 2010. Evolution of variation and variability
624 under fluctuating, stabilizing, and disruptive selection. *Evolution (N Y)* **64**:1912–25.
625 doi:10.1111/j.1558-5646.2010.00979.x
626 Pritchard JK, Pickrell JK, Coop G. 2010. The genetics of human adaptation: Hard sweeps,
627 soft sweeps, and polygenic adaptation. *Curr Biol* **20**:R208–R215.
628 doi:10.1016/J.CUB.2009.11.055
629 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for
630 differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–
631 140. doi:10.1093/bioinformatics/btp616

632 Rocke DM, Durbin B. 2003. Approximate variance-stabilizing transformations for
633 gene-expression microarray data. *Bioinformatics* **19**:966–72.
634 Romero IG, Ruvinsky I, Gilad Y. 2012. Comparative studies of gene expression and the
635 evolution of gene regulation. *Nat Rev Genet* **13**:505–516. doi:10.1038/nrg3229
636 Signor SA, Nuzhdin S V. 2018. The evolution of gene expression in *cis* and *trans*. *Trends*
637 *Genet* **34**:532–544. doi:10.1016/j.tig.2018.03.007
638 Sork VL. 2017. Genomic studies of local adaptation in natural plant populations. *J Hered*
639 **109**:3–15. doi:10.1093/jhered/esx091
640 Suvorov A, Nolte V, Pandey RV, Franssen SU, Futschik A, Schlötterer C. 2013.
641 Intra-specific regulatory variation in *Drosophila pseudoobscura*. *PLoS One* **8**:e83547.
642 doi:10.1371/journal.pone.0083547
643 Turelli M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta
644 meets the abdominal bristle. *Theor Popul Biol* **25**:138–193.
645 doi:10.1016/0040-5809(84)90017-0
646 Wang L, Wang S, Li W. 2012. RSeQC: quality control of RNA-seq experiments.
647 *Bioinformatics* **28**:2184–2185. doi:10.1093/bioinformatics/bts356
648 Wittkopp PJ, Haerum BK, Clark AG. 2008. Regulatory changes underlying expression
649 differences within and between *Drosophila* species. *Nat Genet* **40**:346–350.
650 doi:10.1038/ng.77
651 Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in *cis* and *trans* gene
652 regulation. *Nature* **430**:85–88. doi:10.1038/nature02698
653 Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. 2016. GMAP and GSNAP for
654 genomic sequence alignment: Enhancements to speed, accuracy, and functionality.
655 Humana Press, New York, NY. pp. 283–334. doi:10.1007/978-1-4939-3578-9_15
656 Yoo BH, Nicholas FW, Rathie KA. 1980. Long-term selection for a quantitative character in

large replicate populations of *Drosophila melanogaster*. *Theor Appl Genet* **57**:113–117. doi:10.1007/BF00253881

660 **Figure supplements**

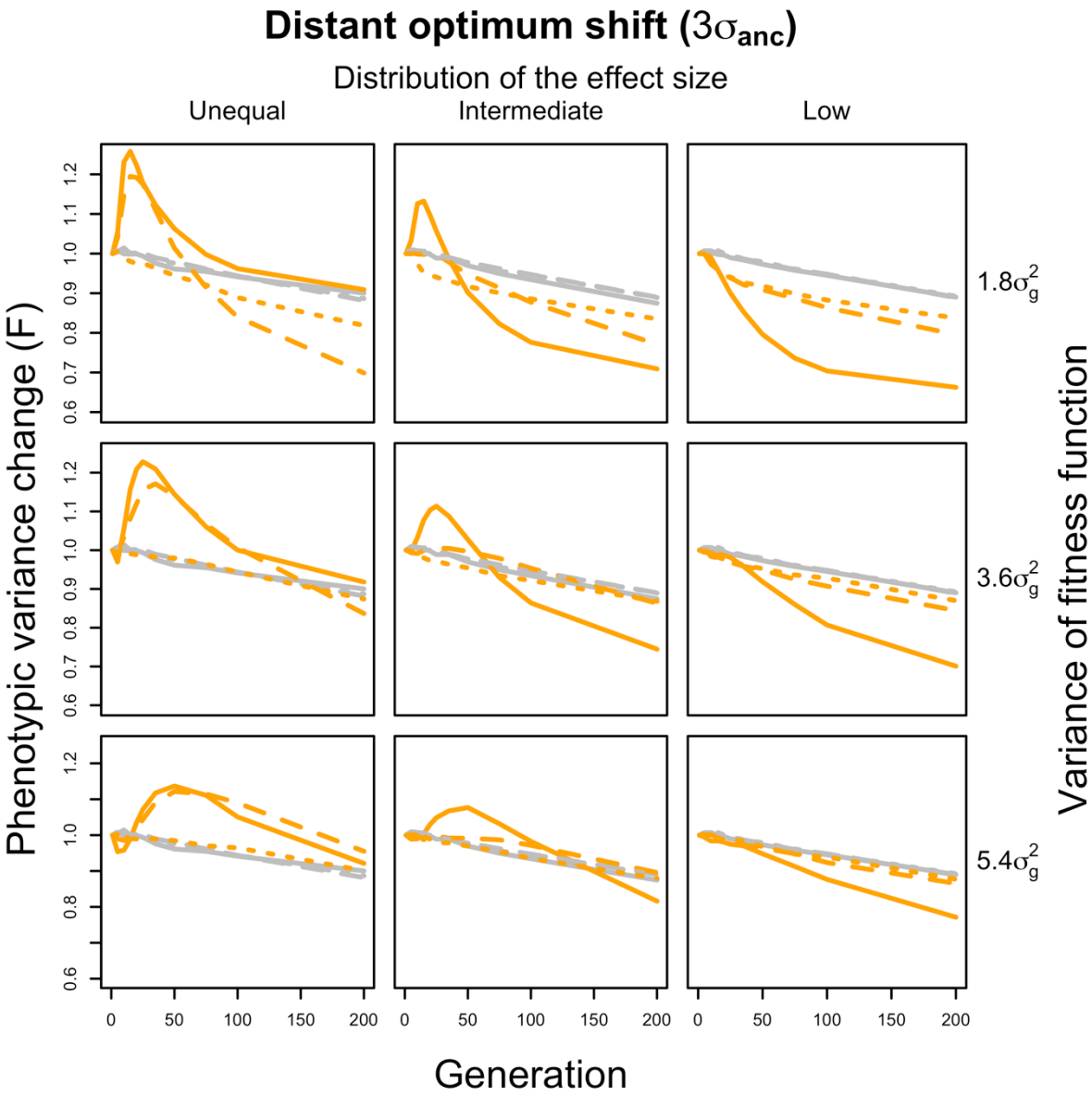


661

662 **Figure 1 – Figure supplement 1. The evolutionary scenario for distant optimum shift.**
663 We consider the case when a quantitative trait (in black) experiences a sudden shift in trait
664 optimum under stabilizing selection. The imposed fitness functions (F.F.) are illustrated in
665 red. The new trait optimum is set away from the ancestral trait mean by three standard
666 deviation of the ancestral trait distribution for distal shift. To vary the strength of stabilizing
667 selection, the variance of the fitness function is set as 1.8, 3.6 and 5.4 standard deviation of
668 the ancestral trait distribution.

669

670



671

672 **Figure 2 – Figure supplement 1. The trajectory of expected changes in phenotypic**
673 **variance when adapting to a distant optimum shift.** The changes in phenotypic variance
674 within 200 generations adapting to a distant optimum shift (orange) are compared to the
675 changes under neutrality (grey) on y axis. The change in variance (F) is calculated as the ratio
676 of phenotypic variance between each evolved time point (generation x) and the ancestral state
677 ($\sigma^2_{200}/\sigma^2_0$). The simulations cover traits controlled by varying numbers of loci underlying the
678 adaptation with three different distributions of effect sizes (columns) and under different

679 strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations have
680 been performed. Only traits with the most (dotted lines, 1000 loci), intermediate (dash lines,
681 50 loci) and the least (solid lines, 5 loci) polygenic architectures are shown. Unlike the
682 continuous decreasing pattern in the cases with moderate optimum shifts, the variance of
683 traits controlled by a few loci (5 loci) with largely dispersed effects would increase first and
684 then decrease when the effect sizes of contributing loci are dispersed (orange solid lines).
685 Nevertheless, for traits with extremely polygenic basis, the phenotypic variance always stays
686 stable over time (orange dotted lines).
687

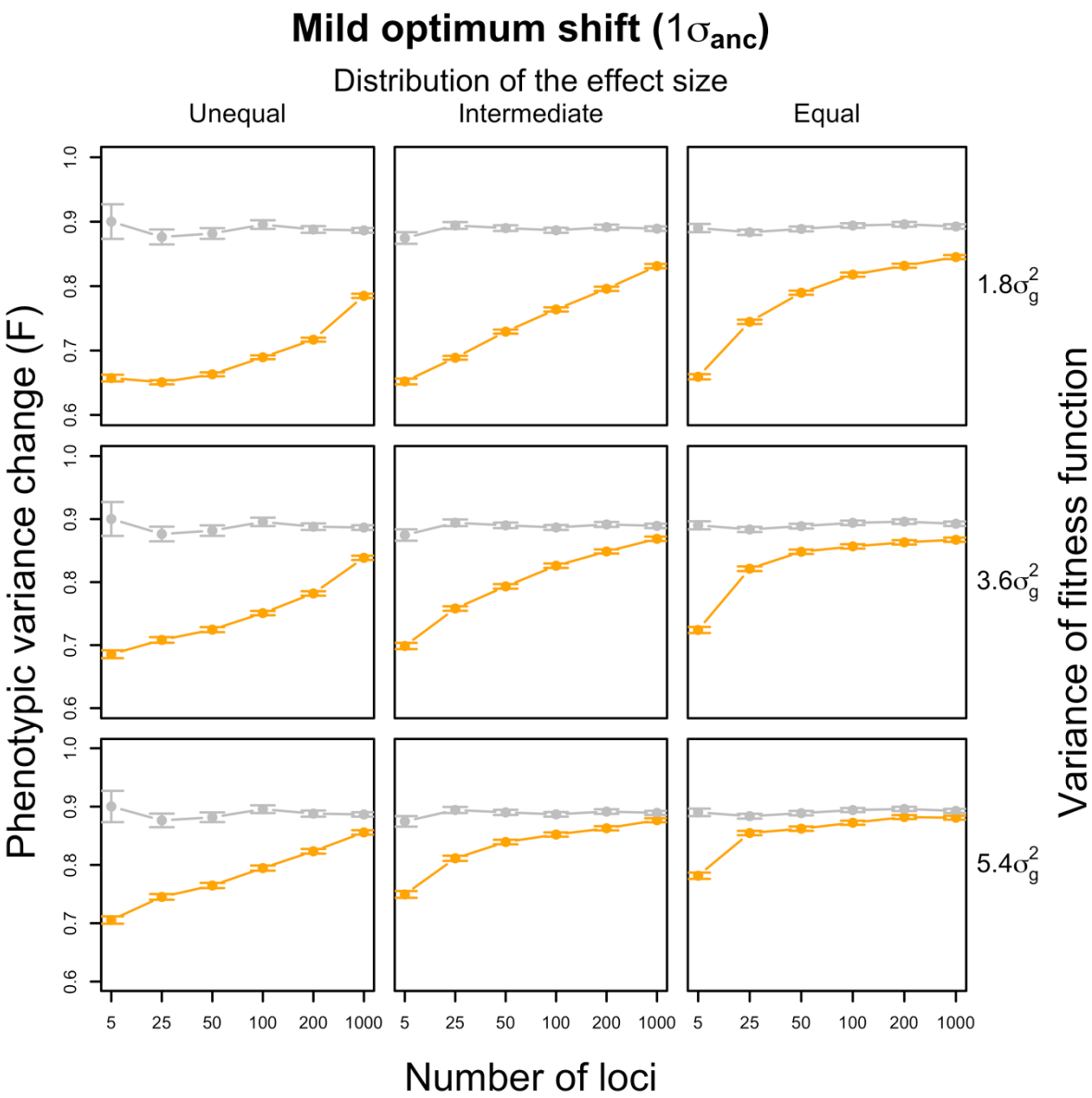


Figure 2 – Figure supplement 2. The expected changes in phenotypic variance and the number of contributing loci. The changes in phenotypic variance after 200 generations adapting to a mild optimum shift (orange) are compared to the changes under neutrality (grey) on y axis. The change in variance (F) is calculated as the ratio between the evolved and ancestral phenotypic variance ($\sigma^2_{200}/\sigma^2_0$). The simulations cover traits controlled by varying numbers of loci underlying the adaptation (x-axes) with three different distributions of effect sizes (columns) and under different strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations have been performed. In all scenarios, the variance of the trait

decreases drastically when the adaptation is controlled by a small number of loci. As the number of contributing loci increases, the phenotypic variance becomes more stable.

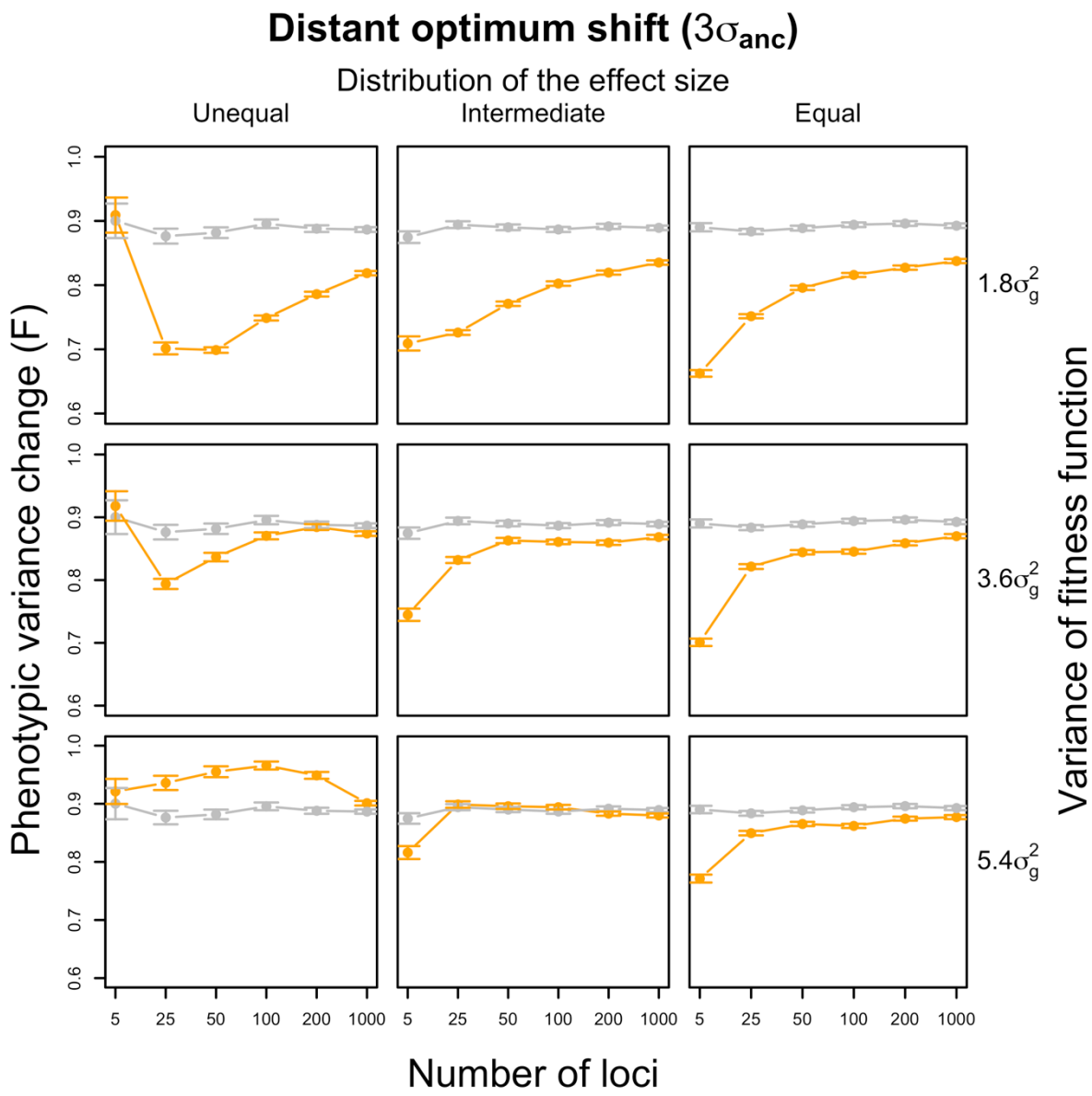
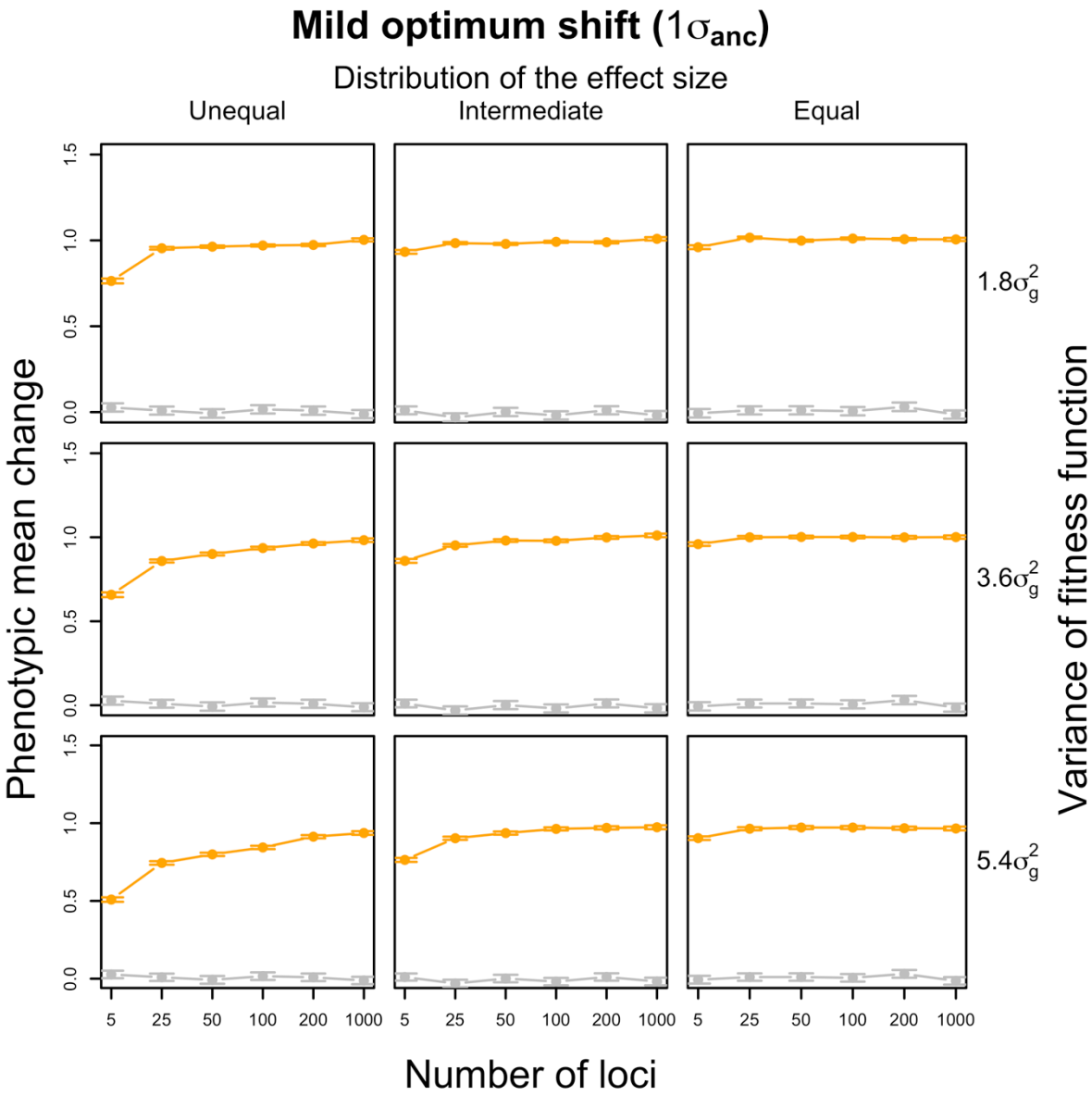


Figure 2 – Figure supplement 3. The expected changes in phenotypic variance and the number of contributing loci. The changes in phenotypic variance after 200 generations adapting to a distant optimum shift (orange) are compared to the changes under neutrality (grey) on y axis. The change in variance (F) is calculated as the ratio between the evolved and ancestral phenotypic variance ($\sigma_{200}^2/\sigma_0^2$). The simulations cover traits controlled by varying numbers of loci underlying the adaptation (x-axes) with three different distributions of effect sizes (columns) and under different strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations have been performed. In most scenarios, the variance of

709 the trait decreases drastically when the adaptation is controlled by a small number of loci. As
710 the number of contributing loci increases, the phenotypic variance becomes more stable.
711 However, exceptions can be observed when the effect sizes of the contributing loci are
712 largely dispersed. In these cases, shift in optimum does not have the strong impact on the
713 variance of traits under simple genetic control (5 contributing loci). In the extreme case,
714 highly dispersed effect sizes in combination with a relaxed phenotypic constraint removes the
715 relationship between the number of loci and the changes in phenotypic variance.
716

717



718

719 **Figure 2 – Figure supplement 4. The changes in phenotypic mean when adapting to a**

720 **mild optimum shift.** The changes in phenotypic mean after 200 generations adapting to a

721 mild optimum shift (orange) are compared to the changes under neutrality (grey) on y axis.

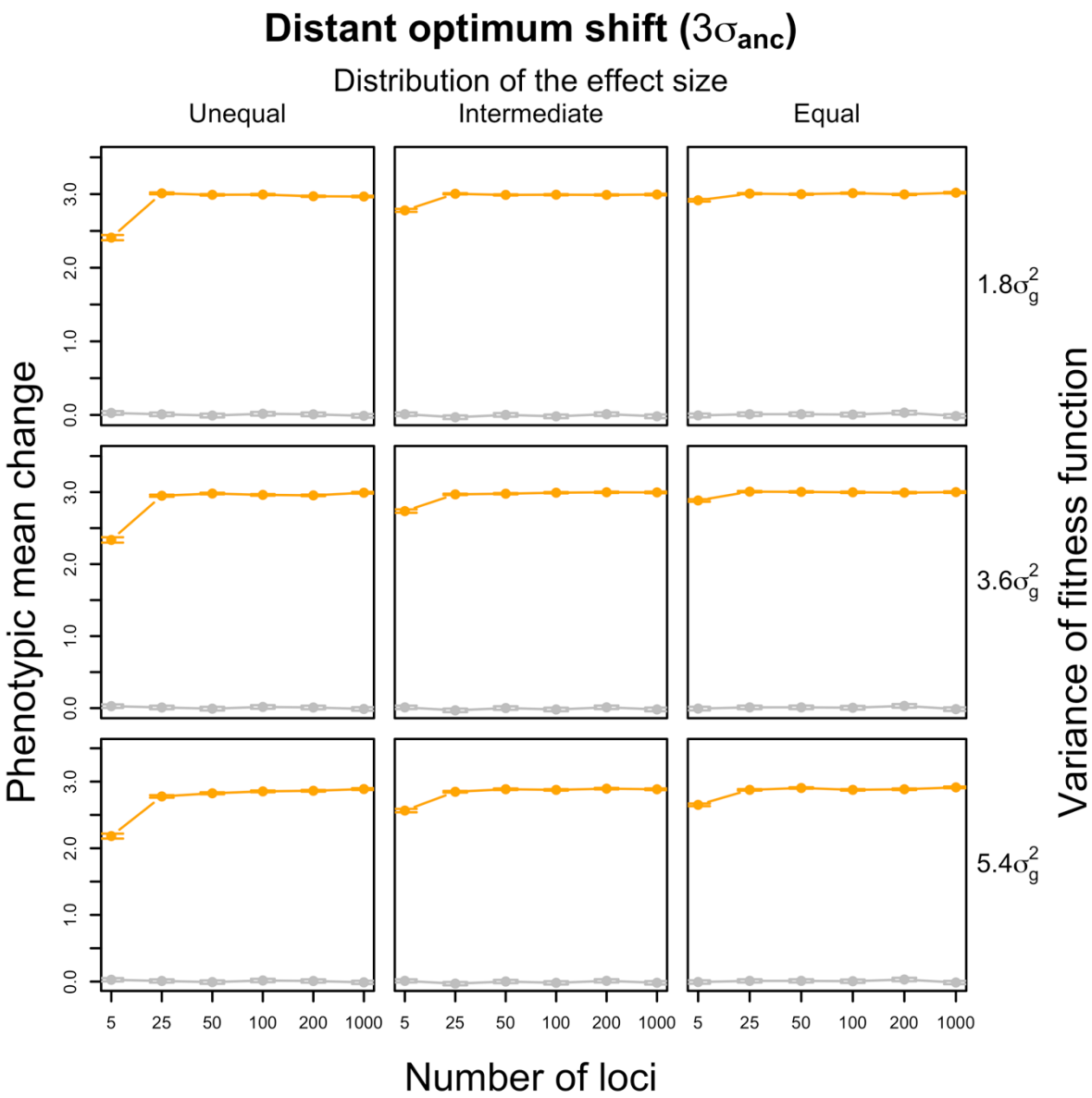
722 The changes in phenotypic mean are scaled by the standard deviation of the ancestral trait

723 distribution. The simulations cover traits controlled by varying numbers of loci underlying

724 the adaptation (x axes) with three different distributions of effect sizes (columns) and under

725 different strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations

have been performed. In most cases, the traits under selection (orange) shift their means by one standard deviation of the ancestral trait distribution (i.e. reaching the new trait optimum) while the neutral traits (grey) stay unchanged.



730

731 **Figure 2 – Figure supplement 5. The changes in phenotypic mean when adapting to a**
732 **distant optimum shift.** The changes in phenotypic mean after 200 generations adapting to a
733 distant optimum shift (orange) are compared to the changes under neutrality (grey) on y axis.
734 The changes in phenotypic mean are scaled by the standard deviation of the ancestral trait
735 distribution. The simulations cover traits controlled by varying numbers of loci underlying
736 the adaptation (x axes) with three different distributions of effect sizes (columns) and under
737 different strength of stabilizing selection (rows). For each scenario, 1000 runs of simulations
738 have been performed. In most cases, the traits under selection (orange) shift their means by

three standard deviation of the ancestral trait distribution (i.e. reaching the new trait optimum) while the neutral traits (grey) stay unchanged.

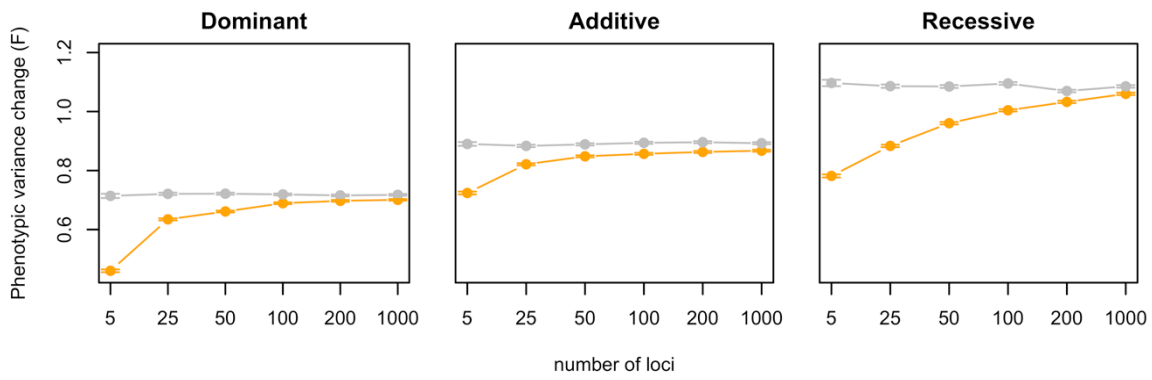
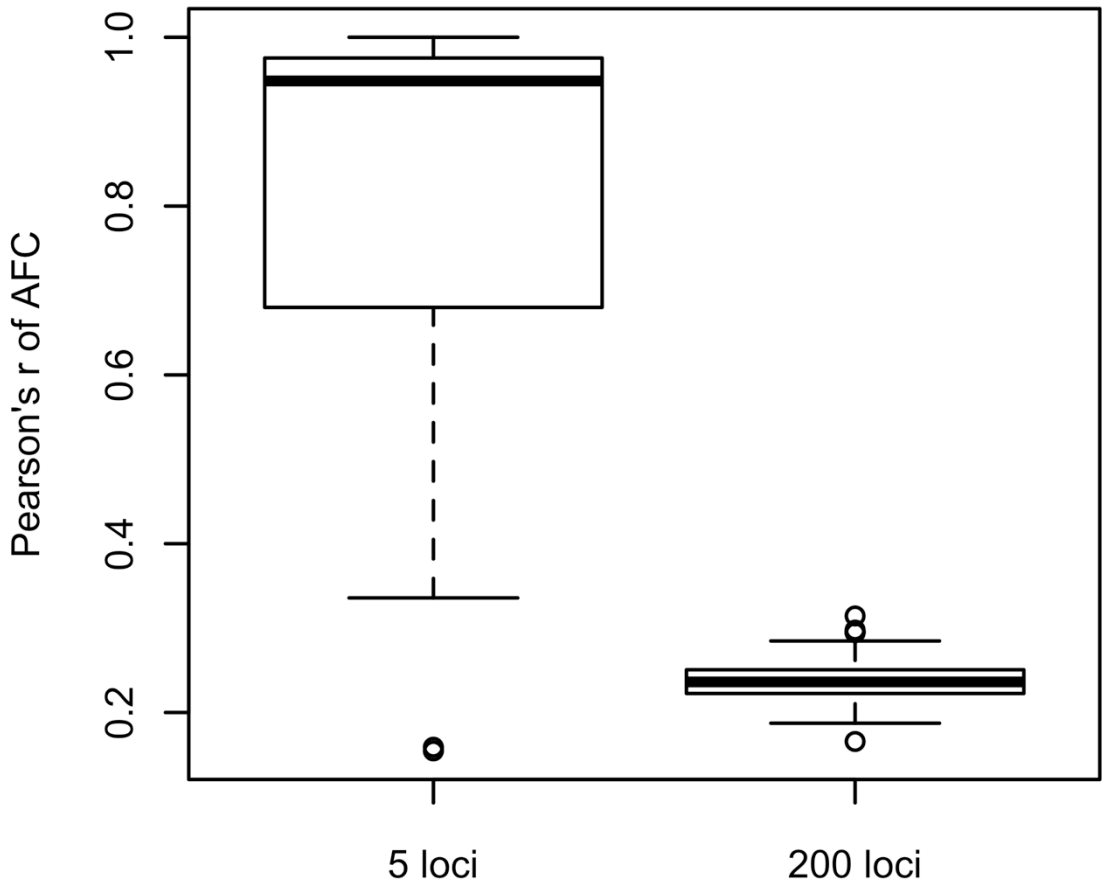


Figure 2 – Figure supplement 6. The expected changes in phenotypic variance for traits controlled by dominant and recessive alleles. The changes in phenotypic variance after 200 generations adapting to a mild optimum shift (orange) are compared to the changes under neutrality (grey) on y axis. The change in variance (F) is calculated as the ratio between the evolved and ancestral phenotypic variance ($\sigma_{200}^2/\sigma_0^2$). This simulation covers traits controlled by varying numbers of loci underlying the adaptation (x axes) with recessive, additive and dominant effects. For each scenario, 1000 runs of simulations have been performed. No matter how the dominance varies, the variance of the trait decreases drastically when the adaptation is controlled by a small number of loci. As the number of contributing loci increases, the phenotypic variance becomes more stable.



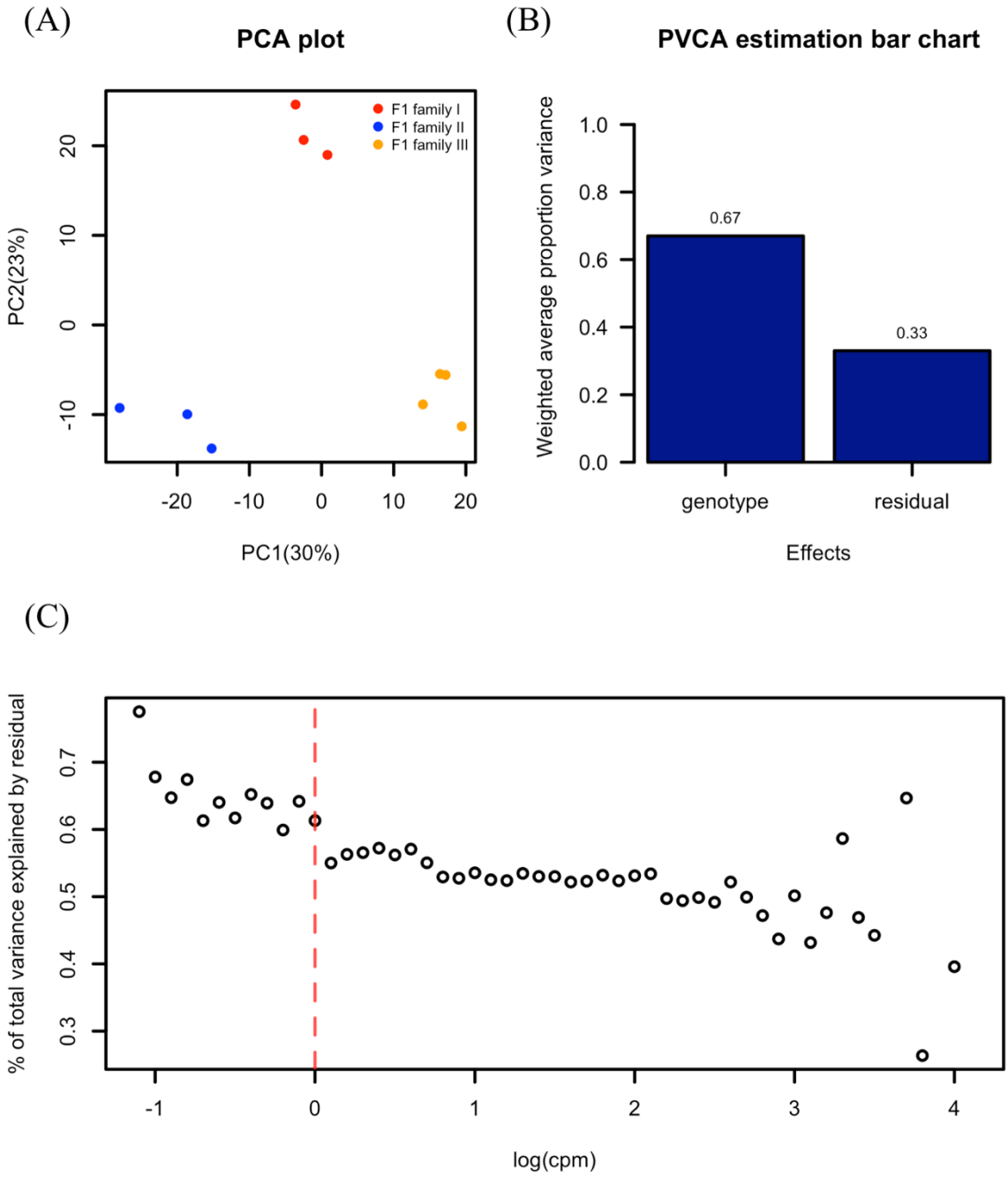
754

755 **Figure 2 – Figure supplement 7. Parallelism in genomic response across 10 evolved**
756 **replicates for traits under the control of five loci and 100 loci.** For the loci with allele
757 frequency change of at least 10% in 200 generations, the average Pearson's correlation
758 coefficient of the frequency change between all pairs of two evolved replicates was
759 calculated to describe the parallelism of the evolution at these loci. An average across loci is
760 used to obtain a general parallelism. 100 runs of simulations with 10 evolution replicates
761 have been performed for each scenario. With five contributing loci, the genomic evolution of

762 the contributing loci is more parallel across the 10 evolved replicates compared to the case

763 when with 100 contributing loci.

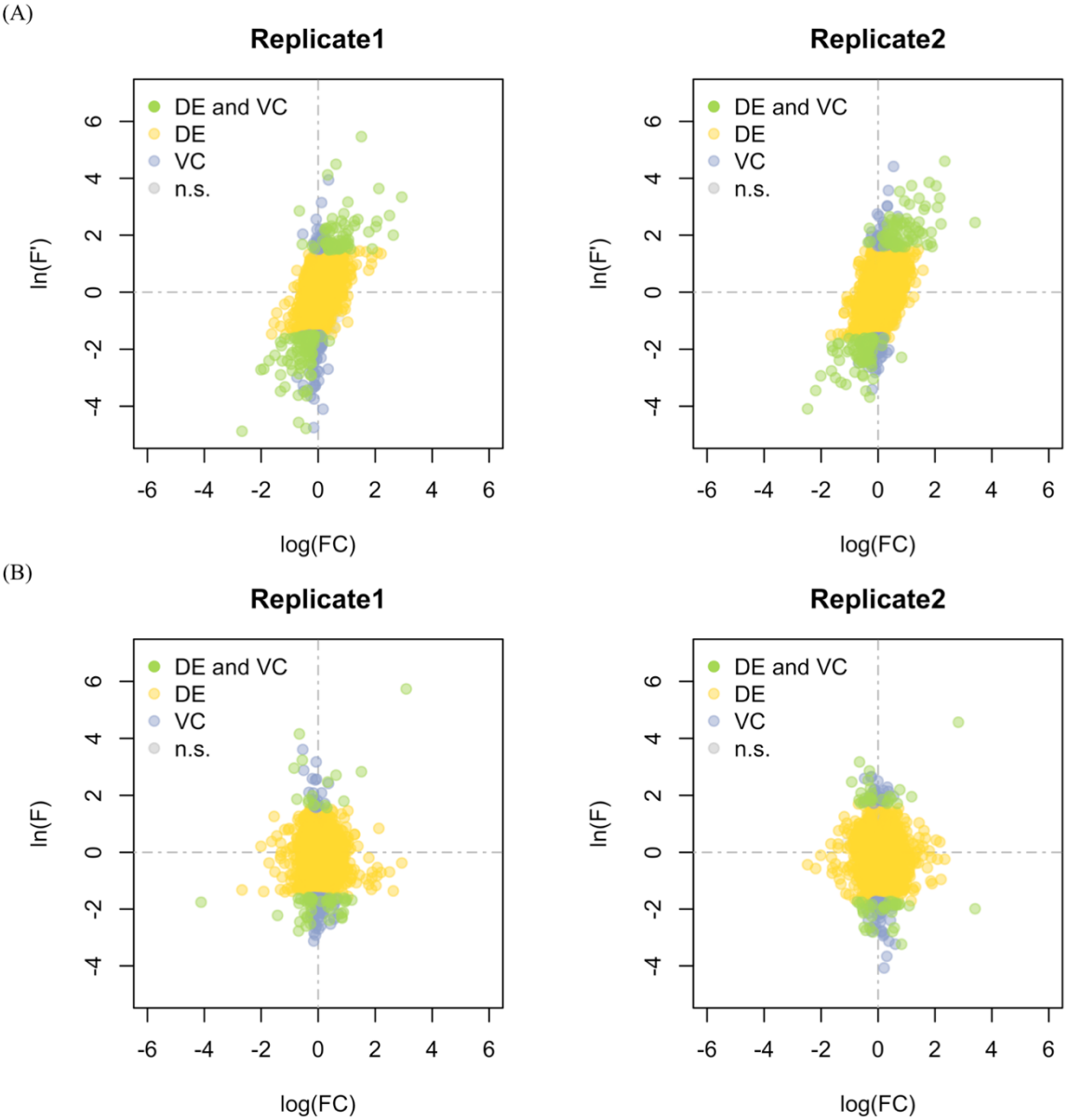
764



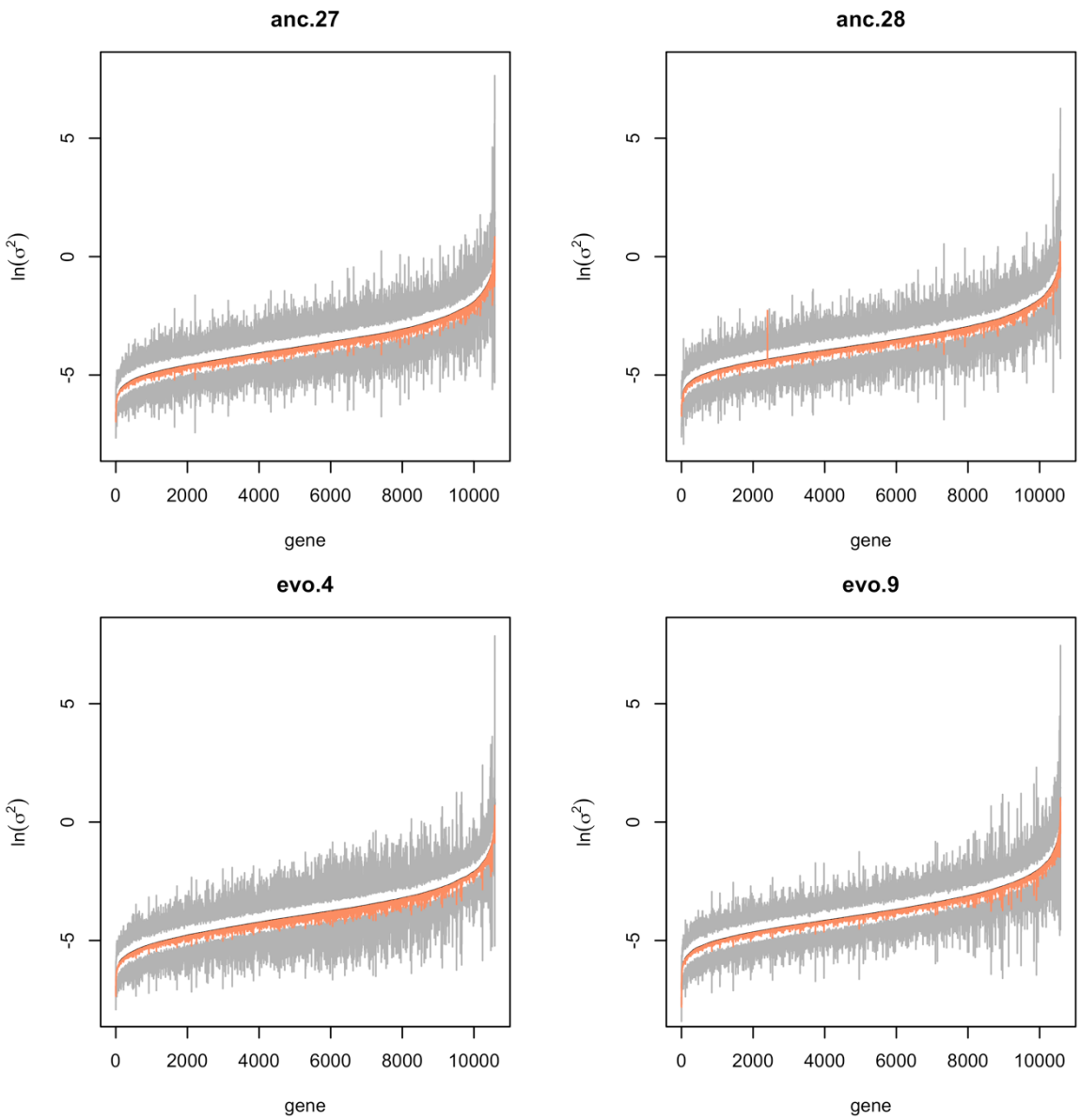
765

766 **Method – Figure supplement 1. Genetic variance in gene expression across F1 families.**
767 **(A)** Principal component analysis (PCA) on the transcriptomes of F1 individuals from three
768 different crosses between the founder iso-female lines. Individuals from different families
769 clustered nicely based on the first two PCs. **(B)** Principal variance component analysis
770 (PVCA) on the transcriptomes of F1 individuals. 67% of the total variance in gene expression
771 was explained by the genetic difference between the individuals. **(C)** Gene-wise analysis of

variance (ANOVA) in gene expression. Genes were binned based on their average expression value (lnCPM) which ranged from -0.8 to 4.1, by bin size of 0.1. The average proportion of variance explained by random error of each bin was visualized. The expression variance of genes with less than 1 count per million bases (CPM) is dominated by residuals.



Method – Figure supplement 2. Log-transformation eliminates the positive relationship between the changes in mean and variance of gene expression. In each panel, the changes in mean expression, log(FC) ($FC = \frac{\overline{y_{evo.}}}{\overline{y_{anc.}}}$) and in variance before **(A)** and after **(B)** the natural log-transformation of each gene were visualized ($F' = \frac{var(y_{evo.})}{var(y_{anc.})}$ and $F = \frac{var(\ln(y_{evo.}))}{var(\ln(y_{anc.}))}$). The positive correlation (r = 0.45) due to the positive mean-variance dependency of negative binomial distribution is removed by the log-transformation (r = -0.05) on gene expression level.



Method – Figure supplement 3. Robustness of the variance estimation using individual sequencing data. Jackknife method was applied to measure the uncertainty of variance estimation. Given a sample size of K, the procedure is to estimate the variance of each gene for K times, each time leaving one sample out. The procedure was conducted independently on 4 populations (anc.27, anc.28, evo.4 and evo.9). In each panel, we visualize Jackknife approximated 95% confidence interval for the variance estimates of each gene. The genes are ordered based on the average variance estimates (black dash line) on the x-axis. The upper and lower limits of the 95% confidence interval are indicated with grey curves. The salmon

line denotes the observed value of the variance estimates. In most cases, the estimates lie in the confidence interval, suggesting robust estimation.

798 **Titles and legend for supplementary files**

799 **Supplementary file 1. Library information of the sample in this study.** This file provides
800 a list of all sequenced samples and the library information.

801

802 **Supplementary file 2. Differential gene expression analysis of two contrasts between**
803 **ancestral and evolved populations.** This file reports the results of DE analysis between anc.
804 27 and evo. 4 (Table S2**A**) and between anc. 28 and evo. 9 (Table S2**B**).

805

806 **Supplementary file 3. F value on the gene expression of two contrasts between ancestral**
807 **and evolved populations.** This file reports the results of gene expression variance
808 comparisons between anc. 27 and evo. 4 (Table S3**A**) and between anc. 28 and evo. 9 (Table
809 S3**B**).

810