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

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, difficult to characterize. Here, we propose to study the adaptive architecture of traits by monitoring the evolution of their phenotypic variance during adaptation to a new environment in well-defined laboratory conditions. Extensive computer simulations show that the evolution of phenotypic variance in a replicated experimental evolution setting can distinguish between oligogenic and polygenic adaptive architectures. We compared 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 suggest that a majority of adaptive gene expression evolution can be explained by a polygenic architecture. We propose that tracking the evolution of phenotypic variance across generations can provide an approach to characterize the adaptive architecture.

Significant Statement

It is widely accepted that most complex traits have a polygenic basis. Nevertheless, it is difficult to predict which of these loci are responding to selection when a population is exposed to a new selection regime. To address this situation, we propose to infer the adaptive architecture for traits by tracking the evolution of their phenotypic variance during adaptation to a new environment. As a case study, we analyze the evolution of gene expression variance in outbred *Drosophila simulans* populations adapting to a new temperature regime to infer the genetic architecture of adaptive gene expression evolution. We suggested that the adaptive gene expression evolution is better explained by a polygenic architecture.

Introduction

It is widely accepted that most complex traits have a polygenic 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 contributing genes are free 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. Characterizing the adaptive architecture by mapping selected loci is not easy, in particular when more than a handful of genes are involved. To circumvent this problem, we introduce an approach, analogous to the Castle-Wright estimator (Castle, 1921), to infer the complexity of the adaptive architecture (i.e. simple with few contributing loci or complex with a polygenic basis). We propose to study the evolution of phenotypic variance, which may provide some insights into the key parameters of the adaptive architecture.

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 & Mackay, 1963). Over the past years, mathematical models have been developed to 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 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 & Turelli, 1987; Keightley & Hill, 1989; Barton & Keightley, 2002; Jain & Stephan, 2015; Franssen *et al.*, 2017; Hayward & Sella, 2019). For instance, for traits with oligogenic architecture, the genetic variance could drop dramatically during adaptation, while with polygenic architectures, only minor effects on the variance are expected (Barton et al., 2017; Franssen et al., 2017; Jain and Stephan , 2015). These insights suggest that a time-resolved analysis of phenotypic variance has the potential to shed light onto the complexity of the underlying adaptive architecture.

Despite its potential importance for the understanding of adaptation, we are faced with the situation that few empirical data are available for the evolution of phenotypic variance. The use of natural populations to study changes in phenotypes, and even more so phenotypic variances, is limited as the environmental heterogeneity cannot be controlled and common garden experiments are required to study the phenotypic variance, which is not feasible for many species. A complementary approach to study the evolution of phenotypic variance is experimental evolution (Kawecki *et al.*, 2012). With replicated populations starting from the same founders and evolving under tightly controlled environmental conditions, experimental evolution provides the opportunity to study the evolution of phenotypic variance.

Most experimental evolution studies in sexual populations focused on the evolution of phenotypic means, rather than variance (e.g.: Chippindale *et al.*, 1996; Burke *et al.*, 2010; Mallard *et al.*, 2018; Jakšić *et al.*, 2020). A notable exception is a study which applied fluctuating, stabilizing and disruptive selection to a small number of wing shape-related traits (Pélabon *et al.*, 2010). Other studies tracked the variance evolution of behavioral and cranial traits under directional selection in mice (Careau *et al.*, 2015; Penna *et al.*, 2017). Instead of looking at a preselected subset of phenotypes which limits the generality, we will focus on gene expression, a set of molecular phenotypes, which can be easily quantified as microarrays and, more recently, RNA-Seq have become available. Importantly, the expression levels of genes exhibit the same properties (e.g.: continuality and normality) as other complex quantitative traits (Mackay *et al.*, 2009). Thus, gene expression has also been widely employed to study the adaptation of locally adapted populations (e.g.: Romero *et al.* 2012; Sork 2017; Signor and Nuzhdin 2018) or ancestral and evolved populations in the context of experimental evolution (e.g.: Lenski *et al.* 1994; Ferea *et al.* 1999; Huang and Agrawal 2016; Mallard *et al.* 2018).

In this study, we performed 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 results that even a moderately polygenic architecture is associated with a high stability of the phenotypic variance of selected traits across different phases of adaptation and evaluate the predictive performance under different scenarios. Applying this insight to a recently published dataset (Lai & Schlötterer, 2022), we show that for putatively selected genes (DE genes), their average changes in expression

variance were indistinguishable from genes without changes in mean expression. We propose that this pattern reflects a polygenic basis of adaptive gene expression evolution.

Results

The central idea of this study is that the complexity of adaptive traits can be inferred 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.: (Jain & Stephan, 2015; Barton *et al.*, 2017; Franssen *et al.*, 2017)), its application to empirical data has been limited. A recent experimental evolution study reporting the phenotypic variance of more than 10,000 gene expression traits before and after 100 generations of adaptation to a novel environment (Lai & Schlötterer, 2022) provides an opportunity to fill this gap. Since gene expression can be viewed as quantitative traits (Ayroles *et al.*, 2009; Hill *et al.*, 2021), we apply the theoretical predictions to the empirical data to infer the genetic basis of adaptive gene expression evolution. Hence, as the first step of this study, we explored to what extent theoretical predictions can be generalized to expression traits obtained from a typical experimental evolution setting by considering a broad parameter space and accounting for linkage. Using population genetic parameters (e.g. number of starting haplotypes, effective population size) estimated from the above-mentioned evolution experiment, we simulated 1000 neutral and 1000 selected expression traits. To account for experimental and environmental noise, we used an empirical distribution of heritabilities (estimated from gene expression measurements across F1 families; see supplementary information). The simulated selection regime consists of a mild/distant shift in trait optimum with weak/intermediate/strong stabilizing selection (Figure 1a and Supplementary Figure S4) (Hayward & Sella, 2019). We assumed additivity and a negative correlation between the ancestral allele frequency and the effect size of contributing loci (Otte et al. 2020) (Figure 1b). With three different distributions of effect size (Figure 1c), we investigated how the number of contributing loci affects the evolution of gene expression variance with and without selection.

We monitored the change in phenotypic variance over 100 generations, which was sufficient to reach the trait optimum for most parameter combinations (Supplementary Figure S5). We compared the change in variance relative to the start of the experiment in populations with and without selection. First, we studied a mild (one standard deviation of the ancestral phenotypic distribution) shift in trait optimum. As expected for a founder population derived from a

substantially larger natural population, we find that even under neutrality, the phenotypic variance does not remain constant, but gradually decreases during 100 generations of experimental evolution (Figure 2). This pattern is unaffected by the genetic architecture of the neutral traits. We explain this loss of variance by the fixation of variants segregating in the founder population and the fact that we did not simulate new mutations, as they do not contribute to adaptation in such short time scales (Burke *et al.*, 2010). Although our simulations used moderate population sizes, they nicely recapitulate the patterns described for populations without drift (Jain & Stephan, 2015; Barton *et al.*, 2017). A pronounced drop in phenotypic variance is observed when a trait is approaching a new optimum with few contributing loci (Figure 2). When more loci (with smaller effects) are contributing to the selected phenotype, the difference to neutrality becomes very small (Figure 2). In addition to the number of contributing loci, the heterogeneity in effect size among loci and the shape of the fitness function have a major impact. The larger the difference in effect size is, the more pronounced was the influence of the number of contributing loci (Figure 2). The opposite effect was seen for the width of the fitness function – a wider fitness function decreased the influence of the number of contributing loci (Figure 2). Importantly, these patterns were not affected by the duration of the experiment - qualitatively identical patterns were seen at different time points until generation 200.

For a more distant trait optimum (three standard deviations of the ancestral phenotypic distribution), we noticed some interesting dynamics that were not apparent for a closer trait optimum (Supplementary Figure S6). The most striking one was the temporal heterogeneity of the phenotypic variance when few loci of unequal effects are contributing. During the early stage of adaptation, the variance increased and dropped later below the variance in the founder population. With an increasing number of contributing loci, this pattern disappeared and closely matched the neutral case (Supplementary Figure S6).

Overall, our simulations indicate that with a larger number of contributing loci the variance fitted the neutral pattern better. Modifying dominance did not change the overall patterns (Supplementary Figure S7). The large influence of key parameters of the adaptive architecture, in particular the number of contributing loci and their effect sizes, on the temporal phenotypic

variance dynamics suggests that it should be possible to exploit this relationship to characterize the genetic basis of phenotypic adaptation.

A potentially interesting application of the relationship between number of contributing loci and evolution of variance is the inference of the adaptive architecture of a given trait. For a single selected trait, a F test contrasting the reduction of variance after adaptation against neutral loss of variance (~0.9 after 100 generations in our simulations) can be used to infer the number of loci contributing to adaptive evolution. The lower bound for the number of contributing loci can be determined by the largest number of loci that results in a significant difference between a selected trait and the neutral expectation in all parameter combinations. Nevertheless, the power of this approach depends strongly on the sample size. When the entire population (N = 300) is phenotyped for the focal trait, the power to rule out an adaptive architecture of no more than 5 loci by significant decrease in variance of a selected trait is only 44%. This indicates the limited ability to distinguish polygenic from oligogenic architectures of single traits. With a more realistic sample size of 20 individuals the distinction is even more difficult such that we conclude that it is not possible to infer the adaptive architecture for a single trait when only two time points are available (Supplementary Figure S8).

Alternatively, it is possible to study multiple, presumably independent, selected phenotypes together. Assuming a similar genetic architecture of the selected traits, we showed that the distribution of variance changes of 1000 selected traits with a simple genetic architecture (five loci) can be distinguished from neutral changes. Even with a sample size of 20 individuals, as in our experimental data (see Materials and Methods), significant differences from neutral expectations can be detected for some parameters with a power close to 100% (Figure 3 and Supplementary Figure S9). This suggests that oligogenic and polygenic architectures of a group of selected traits can be distinguished experimentally even with moderate sample sizes.

In the simulations, we fixed most experimental parameters such as population sizes, generation and sample sizes to reflect the experimental design of typical experimental evolution studies. Nevertheless, for future experimental evolution studies it is important to understand how the experimental design affects the power. We performed additional simulations to explore the

196 impact of the sample size, number of generations and population size. As expected, the power
197 increases with larger sample size, allowing to discriminate architectures with a larger number
198 of contributing loci (Figure 4a). Later generations are also more informative than earlier
199 generations (i.e., higher power to reject a larger number of loci) (Figure 4b). Larger population
200 sizes could also improve the determination of polygenicity, as we find an increase in power for
201 a population size of 1200 compared to 300 (Figure 4c). In summary, larger and longer evolution
202 experiments with larger sample size are superior. They provide a better discrimination of
203 genetic architectures with a larger number of contributing loci-i.e. a better distinction between
204 oligogenic and polygenic architectures.

205

206 As a case study, we investigated the evolution of gene expression variance in replicated
207 populations evolving in a new hot temperature regime and inferred the adaptive architecture of
208 gene expression traits. The evolved populations were derived from the same ancestral
209 population but evolved independently for more than 100 generations in a novel temperature
210 regime with daily temperature fluctuations between 18 and 28°C (Figure 5a). Rather than
211 relying on pooled samples which only allow mean estimates, we quantified gene expression of
212 19-22 individuals from reconstituted ancestral populations (F0) and two evolved populations
213 (F103) in a common garden setup (Lai & Schlötterer, 2021). Principal Component Analysis
214 (PCA) indicated that 11.9% of the variation in gene expression can be explained by the first
215 PC which separates evolved and ancestral populations, reflecting clear adaptive gene
216 expression changes in response to the novel hot temperature regime (Figure 5b). The means
217 and variances of the expression of each gene were estimated and compared between the two
218 reconstituted ancestral populations and the two evolved populations separately (see Materials
219 and methods). Due to the usage of different lot numbers for the RNA-Seq library preparation
220 (Table S1), we only contrasted ancestral and evolved samples generated with the same lot
221 number (see Materials and methods) to avoid any unnecessary confounding effects.

222

223 The comparison of ancestral and evolved populations identified 2,775 genes in the first
224 population and 2,677 genes in the second population which significantly changed mean
225 expression in the evolved flies (FDR<0.05, (Lai & Schlötterer, 2022)). 1,172 differentially
226 expressed genes are shared between two populations (Figure 6a). The concordance of both
227 populations and the similarity to previous studies of the same evolution experiment (Hsu *et al.*,

228 2020; Jakšić *et al.*, 2020; Lai & Schlötterer, 2022) suggests that most of the altered expression
229 means are mainly driven by selection, rather than by drift. We scaled the gene expression
230 change with the standard deviation in the ancestral population to approximate the selection
231 strength on each gene. The differentially expressed genes in both populations showed a broad
232 distribution of expression change, but the averaged mean expression changed by one standard
233 deviation (Figure 6b), which is unlikely to be reached by drift based on our simulations (less
234 than 1% of the neutral traits drifted by one standard deviation of the ancestral phenotypic
235 variance). Assuming that these putative adaptive gene expression phenotypes reached the new
236 trait optimum, this corresponds, on average, to a mild shift in trait optimum in our computer
237 simulations. Reasoning that genes with significant mean expression changes are under
238 selection and the rest of the transcriptome is not responding to the new environment (neutral),
239 we compared the evolutionary changes in expression variance of genes with and without
240 significant changes in expression mean (DE and non-DE genes), to characterize the adaptive
241 architecture of gene expression. The underlying hypothesis is that a simple adaptive
242 architecture will be associated with a significant drop in the expression variance of the selected
243 genes, while for polygenic gene expression evolution, the variance remains similar to neutral
244 genes. Our computer simulations showed that this expectation from theoretical results (Jain &
245 Stephan, 2015; Barton *et al.*, 2017; Franssen *et al.*, 2017) also holds for moderately sized
246 populations and linked loci (Figure 2).

247

248 The reduction in population size from a natural population to laboratory cultures would result
249 in a continuous reduction in phenotypic variance over time even under neutrality, as shown in
250 the simulations (Figure 2). Hence, a robust inference of the adaptive architecture of gene
251 expression relies on the comparison of putative selected (DE) and neutral (non-DE) genes. In
252 both replicates the expression variance of putative neutral (non-DE) genes dropped to 84% of
253 the ancestral variance, which is quite close to the neutral cases in our simulations. While it
254 would be desirable to test each putatively selected (DE) gene independently, the moderate
255 sample size (~20 per population) does not provide sufficient power to do so (Supplementary
256 Figure S8). Rather, we considered all putatively selected (DE) genes jointly and compared their
257 variance changes to the ones of putatively neutral (non-DE) genes. Remarkably, we find that
258 the changes in variance of putatively selected genes with significant mean expression changes
259 are indistinguishable from the genes that do not change their mean expression (t test, p-value
260 > 0.05; Figure 6c). The magnitude of gene expression variance changes in both sets of genes

261 (median F-value = 0.85 and 0.84 for DE and non-DE genes, respectively, in population 1, and
262 0.82 and 0.84 in population 2) is quite close to the neutral conditions in our computer
263 simulations. This suggests that selection on mean expression did not significantly change the
264 expression variance during adaptation. This is further supported by the lack of correlation
265 between the evolutionary changes in gene expression variance and mean (Lai & Schlötterer,
266 2022). In our simulations we assumed that the traits are independent from each other, which
267 may not be the case for gene expression data. Co-regulation (non-independence) among gene
268 expression may affect our inference. We address this issue by using only 1000 genes with
269 reduced correlation to diminish the impact of gene co-regulation for the analysis (see
270 supplementary information and supplementary figure S10). Despite this down-sampling, we
271 still observed no difference in variance change between DE and non-DE genes using this
272 approach (Supplementary figure S11). Based on our analyses we suggest that variation in
273 expression of most putatively adaptive genes has polygenic basis (i.e. more than five
274 contributing loci).

275

276 Since we only explored two time points (F0 and F103) rather than a full time series, it may be
277 possible that an oligogenic basis could also result in a similar phenotypic variance as neutral
278 phenotypes like what is observed for a polygenic architecture (Supplementary Figure S9). This
279 can be seen in an intuitive case when a single/few major effect allele(s) starts at a low frequency
280 and becomes fixed (Yoo *et al.*, 1980). Since an oligogenic basis results in a highly parallel
281 genomic selection response (Supplementary Figure S12), it is possible to distinguish polygenic
282 and oligogenic architectures with phenotypes from two time points only when genomic data
283 are available. Because the genomic signature in the same experiment uncovered a highly
284 heterogeneous selection response (Barghi *et al.*, 2019), we can exclude the unlikely explanation
285 of an oligogenic architecture resulting in a similar expression variance as non-selected genes.
286 Rather, we suggest that the adaptive response in gene expression could be explained by a
287 polygenic architecture (i.e. more than five contributing loci).

288

289

290 **Discussion**

291 Population genetics has a long tradition of characterizing adaptation based on the genomic
292 signature of selected loci (Nielsen, 2005). Nevertheless, for selected phenotypes with a
293 polygenic architecture, the contribution of individual loci to phenotypic change may be too
294 subtle to be detected with classic population genetic methods (Pritchard *et al.*, 2010). Although
295 identification of many small-effect loci is challenging, it may be possible to determine the
296 number of contributing loci based on the dynamics of phenotypic variance.

297

298 Inspired by the Castle-Wright estimator (Castle, 1921) that estimates the number of loci
299 contributing to a quantitative trait from the phenotypic variance of the F2, we propose that the
300 temporal heterogeneity of the phenotypic variance can potentially be used to infer the number
301 of loci contributing to the adaptive response of a phenotype as well as other parameters of the
302 adaptive architecture. Reasoning that experimental evolution is probably the best approach to
303 obtain phenotypic time series, we performed computer simulations specifically tailored to
304 typical experimental evolution studies with *Drosophila*. We demonstrated that, in an
305 experimental evolution setup, the temporal dynamics of the phenotypic variance is strongly
306 affected by the number of contributing loci and other parameters of the adaptive architecture
307 such as the distribution of effect size.

308

309 We acknowledge that the proposed approach for the characterization adaptive architecture
310 critically depends on the power to distinguish the variance changes of selected traits from the
311 neutral expectations. Since the accurate inference of gene expression variance is key, we
312 partitioned the expression variance of each gene among three independent F1 families consist
313 of almost identical, heterozygous individuals (see Supplementary Information). We show that
314 on average, ~50% of expression variance can be attributed to random error - a combination of
315 technical (library preparation, RNA-extraction, sequencing) and biological (stochastic gene
316 expression differences among individuals) noise. We included this error in our simulations by
317 adopting the empirical heritability distribution (heritability of every gene inferred from this
318 experiment) such that random errors relative to the ancestral phenotypic variance were
319 introduced to the calculation of phenotypic values at each generation for each simulated trait.
320 The heritability estimates in our experiment may be too conservative (i.e. too high) because we
321 assumed no environmental heterogeneity. We also used heritability estimates from another

322 gene expression study on *Drosophila* which included more lines and accounted for
323 environmental heterogeneity (Ayroles et al., 2009). The results are qualitatively similar (Figure
324 3).

325

326 Another potential limitation of our analysis is that we assumed neutrality of non-DE genes, it
327 may, however, be possible that they are also subject to stabilizing selection, but without shift
328 in trait optimum. In this case, the same change in variance is expected for both groups of traits
329 - even under an oligogenic architecture, as shown by computer simulations (Supplementary
330 Figure S13a and b, see supplementary information). If stabilizing selection is also operating on
331 non-DE genes, the change in variance should be correlated between replicate populations
332 because each trait experiences the same strength of stabilizing selection (each trait has the same
333 heritability in both replicates, but it differs between traits). This prediction was confirmed in
334 our computer simulations (Supplementary Figure S13c). In our empirical data, however, the
335 variance change was more strongly correlated between two evolution replicates for DE genes
336 ($r=0.07$) than for non-DE genes ($r=0.02$) differ from each other, implying less selection for
337 variance on non-DE genes. This difference in correlation is highly significant when we
338 compared the correlation coefficient of DE genes to the same number of randomly drawn non-
339 DE genes ($p<0.01$, Supplementary Figure S13d). Given that the distribution of correlation
340 coefficients for non-DE genes is very close to the expectations under neutrality from computer
341 simulations, we think that the behavior is better approximated by neutrality for most of the
342 non-DE genes, rather than assuming that similar levels of stabilizing selection for both classes
343 of genes. It is not clear, however, whether our results reflect a much broader fitness function
344 determining the evolution of non-DE genes per se, or the simple laboratory environment
345 relaxes selection on non-DE genes.

346

347 Given the limited power to infer the number of contributing loci for each expression phenotype
348 (Supplementary figure S8), we grouped all putatively selected expression traits and limited our
349 inference of the adaptive architecture to these genes jointly. This approach makes the implicit
350 assumption that all expression traits in this group are independent of each other and have
351 similar level of complexity in their adaptive architectures. Thus, the joint inference could
352 potentially compromise the detection of a minor subset of genes whose expression evolution
353 is under simple genetic control.

354

355 Our results may also provide some insights into the molecular basis of gene expression
356 evolution. Gene expression is modified by many trans-acting factors and some cis-regulatory
357 variation (Hill *et al.*, 2021). Adaptive gene expression changes can be either driven by
358 polymorphism in *cis*-regulatory elements or by *trans*-acting variants. While interspecific
359 differences in gene expression are predominantly caused by *cis*-regulation (Wittkopp *et al.*,
360 2004), intraspecific variation is mostly driven by trans regulatory changes (Wittkopp *et al.*,
361 2008; Suvorov *et al.*, 2013). Adaptive gene expression changes which are well-characterized
362 on the molecular level typically have a *cis*-regulatory basis that is not only frequently
363 associated with the insertion of a transposable element (e.g.: (Daborn *et al.*, 2002)), but also
364 sometimes with multiple regulatory variants (Endler *et al.*, 2018). Two lines of evidence
365 suggest that *cis*-regulatory variation cannot be the driver of adaptive gene expression changes
366 observed in this study. First, the mutational target size is too small to harbor a sufficiently large
367 number of alleles segregating in the founder population. Second, too few recombination events
368 occur during the experiment to uncouple regulatory variants located on a given haplotype such
369 that they could generate a signal of polygenic adaptation. More likely, the polygenic adaptive
370 architecture of gene expression change reflects the joint effects of many *trans*-acting variants.

371

372 Because we could only analyze phenotypic data from two time points, the founder population
373 and replicate populations evolved for 103 generations, we were not able to obtain a more
374 quantitative estimate of the number of contributing loci, in particular as other parameters of the
375 adaptive architecture are not known and need to be co-estimated. In addition, with only two
376 time points, for a few parameter combinations, an oligogenic response can also result in a
377 similar phenotypic variance change as a polygenic one but with a much higher parallel response
378 of genomic markers (Supplementary Figure S12). Hence, not only more time points describing
379 the phenotypic trajectory, but also some genomic data could contribute to inferring the adaptive
380 architecture in experimental evolution studies.

381

382 The extension of the concept proposed in this study to natural populations could face several
383 challenges that warrant extra caution and further investigation. First, the above-mentioned
384 challenge of sample size would be more pronounced in a natural setup. Second, phenotypic
385 time series over evolutionary relevant time scales are costly (but see (Clutton-Brock &

386 Pemberton, 2004)) and third, the distinction of environmental heterogeneity from genetic
387 changes is considerably more challenging than under controlled laboratory conditions.

388

389 **Materials and Methods**

390 Partitioning gene expression trait variance in three F1 families

391 Since the accurate measurement of gene expression variance is key to this study, we analyzed
392 the expression variance for each gene among three independent F1 families obtained from
393 crossing two highly inbred strains (see Supplementary Information and Table S1). We show
394 that on average, ~50% of expression variance can be attributed to random error - a combination
395 of technical (library preparation, RNA-extraction, sequencing) and biological (stochastic gene
396 expression differences among individuals) noise (Supplementary Figure S1). This confirms the
397 importance to separate the biological and technical variance in the dataset, as we did in the
398 empirical data analysis (see below).

399

400 Computer simulations

401 We performed forward simulations with MimicrEE2 (Vlachos & Kofler, 2018) using the qff
402 mode to illustrate the influence of the genetic architecture on the evolution of phenotypic
403 variance during the adaptation to a new trait optimum (Figure 1a). With 189 founder haplotypes
404 (Barghi *et al.*, 2019), we simulated quantitative gene expression traits under the control of
405 different numbers of loci (M = 5, 25, 50, 100, 200 and 1000) in a population with an effective
406 population size of 300 (reflecting the estimated effective population size in an experimentally
407 evolving population (Barghi *et al.*, 2019)). We used the empirical linkage map from *Drosophila*
408 *simulans* (Howie *et al.*, 2019) to account for linkage. For each trait, we assume an additive
409 model and a negative correlation (r = -0.7, reflecting the observed relationship from Otte *et al.*
410 (2020)) between the effect size and starting frequency (Figure 1b). We used the *correlate()*
411 function implemented in "fabricatr" R package (Blair *et al.*, 2019) to generate the effect sizes.
412 The distribution of the effect sizes of loci controlling a trait can vary depending on the shape
413 parameter of gamma sampling process (shape = 0.5, 2.5 and 100, Figure 1c). The mean of the
414 effect sizes from 3 different shape parameters was standardized to 1/M. As the focus of our
415 study is on expression traits, we used the distribution of heritabilities obtained from empirical
416 gene expression data (Ayroles et al. 2009 and this study, see supplementary information) to
417 assign the environmental/technical variance to a simulated trait (see supplementary
418 information; Supplementary Figure S1) (i.e. for each simulated trait, its heritability and
419 associated environmental/technical variance are sampled from the empirical distribution). We
420 do not simulate de novo mutations as they are not contributing to adaptation on this short time

421 scale (Burke *et al.*, 2010). To simulate stabilizing selection with trait optimum shift, we
422 provided the Gaussian fitness functions with mean of $\overline{X_{anc.}} + a\sqrt{V_{anc.}}$ and standard deviation
423 of $b\sqrt{V_{anc.}}$, where $\overline{X_{anc.}}$ is the ancestral phenotypic mean and $V_{anc.}$ is the ancestral genetic
424 variance (Figure 1a). Parameter "a" determines the distance of optimum shift, which is set to
425 one (similar to the empirical case, Figure 6b) or three (following (Hayward & Sella, 2019)).
426 Parameter "b" indicates the phenotypic constraint would be at trait optimum. The value 3.6 for
427 parameter "b" was taken from Hayward and Sella (2019). In this study, we increase and
428 decrease it by 50% to explore its impact (1.8 or 5.4). For the neutral case, we assumed no
429 variation in fitness among all individuals. For each scenario, 1000 traits were simulated, and
430 each trait was affected by a different set of loci. We compared the phenotypic variance of
431 selected and neutral traits for each genetic architecture. For each trait under each scenario, the
432 phenotypic variance was estimated at different generations and compared to the ancestral
433 phenotypic variance at generation 1 to illustrate the dynamic of phenotypic variance during the
434 evolution ($F = \sigma_x^2/\sigma_1^2$), where x stands for the number of generations. We note that we do not
435 assume that the ancestral population has reached an equilibrium because the ancestral
436 population in a typical experimental evolution study is often phenotyped in the new
437 environment.

438 In addition, we evaluated how different experimental designs (e.g.: different population sizes
439 in the experiment) affect the ability to discriminate between different adaptive architectures.
440 We simulated the evolution of traits controlled by a different number of loci (M = 5, 25, 50,
441 100, 200 and 1000) with varying effect sizes (shape parameter of gamma sampling process =
442 2.5) using a Gaussian fitness function with a mean of $\overline{X_{anc.}} + \sqrt{V_{anc.}}$ and a standard deviation
443 of $3.6\sqrt{V_{anc.}}$ with population size = 1200.

444

445 <u>Evaluating the ability to distinguish adaptive architectures based on trait variance evolution</u>

446 We propose that the temporal dynamics of the phenotypic variance of selected traits can be
447 exploited to characterize the genetic basis of adaptation. This concept could be applied either
448 to a single trait under selection or to a group of selected traits. Depending on whether a single
449 trait or a group of traits is studied, different research objectives can be pursued. Generally
450 speaking, a more complex architecture is expected to result in a variance pattern that resembles
451 neutral evolution (i.e. no change in variance). A lower bound for the number of contributing

452 loci can be inferred by asking for the maximum number of contributing loci causing a variance
453 change during evolution, which significantly deviates from neutral expectations. With
454 increasing uncertainty about the key parameters, the precision of this approach is reduced.
455 Given that many parameters are typically not known, we performed computer simulations for
456 different numbers of loci with a distribution of effect sizes and asked for each focal parameter
457 combination whether a significant deviation from neutral expectations was observed. The
458 number of simulation runs, which deviate from neutral expectations for *x* loci is used as power
459 estimate to reject the null hypothesis of no difference to the variance under neutrality under a
460 given architecture and sample size. In other words, it reflects the confidence to exclude an
461 adaptive architecture with *x* or fewer contributing loci when we observed no significant
462 difference in variance.

463

464 *Power estimates for the analysis of a single trait:* For each simulated trait under different
465 genetic control, we tested whether the variance of this trait changed more than expected under
466 neutrality (F of 0.9 according to the neutral simulations) after 100 generations of selection.
467 First, we assumed that all simulated 300 individuals were phenotyped at both time points. For
468 the 1000 traits with a given genetic architecture (*x* contributing loci) under each selection
469 scenario, we calculated the power to reject the null hypothesis of an adaptive architecture no
470 difference to the variance under neutrality under a given architecture and sample size. Since it
471 is not possible to phenotype all individuals, we also investigated whether and how a reduced
472 sample size compromises the power to reject an adaptive architecture of a single trait based on
473 the change in variance. We randomly selected 20 individuals from each simulation run to
474 estimate the variance and tested whether the magnitude of variance change for a given trait
475 under selection is significantly different from the neutral expectation.

476

477 *Power estimates for the analysis on a group of traits:* We assessed the power to infer the
478 adaptive architecture for a group of selected traits under the assumption that most traits of
479 interest have similar level of complexity. In this case, a test on the average variance changes
480 (F) between a group of selected traits and another group of neutral traits is required. We
481 simulated 1,000 independent selected and 1,000 neutral traits for each genetic architecture and
482 selection regime, and estimated the variance before and after 100 generations of evolution for
483 the selected and neutral traits. For each trait, we measured the phenotypic variance for a sample

484 of 20 individuals and we performed a t-test to compare the distribution of variance change (F)
485 between selected and neutral traits. The power was determined by 100 iterations of sampling
486 20 individuals from a simulated population of 300 individuals.

487 We further evaluated how different experimental designs affect the power to detect simpler
488 adaptive architectures. We used the data from the scenario with shape parameter 2.5 and
489 Gaussian fitness functions with mean of $\overline{X_{anc.}} + \sqrt{V_{anc.}}$ and standard deviation of $3.6\sqrt{V_{anc.}}$.
490 With this dataset, we calculated the power for different sample sizes (n=20, 30, 50, 70),
491 generation times (gen=25, 50, 100, 200) and population sizes (N=300 and 1200).

492

493 Experimental evolution

494 The setup of populations and evolution experiment have been described by Barghi *et al.* (2019).
495 Briefly, ten outbred *Drosophila simulans* populations seeded from 202 isofemale lines were
496 exposed to a laboratory experiment at 28/18 °C with 12hr light/12hr dark photoperiod for more
497 than 100 generations. Each replicate consisted of 1000 to 1250 adults for each generation.

498

499 Common garden experiment

500 The collection of samples from the evolution experiment for RNA-Seq was preceded by two
501 generations of common garden (CGE). The common garden experiment was performed at
502 generation 103 of the evolution in the hot environment and this CGE has been described in
503 (Barghi *et al.*, 2019; Hsu *et al.*, 2019, 2020; Jakšić *et al.*, 2020; Lai & Schlötterer, 2022). In
504 brief, an ancestral population was reconstituted by pooling five mated females from 184
505 founder isofemale lines (Nouhaud *et al.*, 2016). No significant allele frequency differences are
506 expected between the reconstituted ancestral populations and the original ancestral populations
507 initiating the experiment (Nouhaud *et al.*, 2016). Furthermore, we do not anticipate that
508 deleterious alleles acquired during the maintenance of the isofemale lines had a major impact
509 on the phenotypic variance in the reconstituted ancestral population. The reason is that novel
510 deleterious mutations occurring during the maintenance of the isofemale lines are present in a
511 single isofemale line only. Given the large number of isofemale lines (184), such deleterious
512 alleles occur in a low frequency in the reconstituted population with a small influence on the
513 phenotypic variance (Walsh & Lynch, 2018). Furthermore, most of these deleterious alleles
514 are present in heterozygous individuals and masked because deleterious alleles tend to be

515 recessive (Charlesworth & Charlesworth, 2010). As described previously (Lai & Schlötterer,
516 2022), two replicates of the reconstituted ancestral population and two independently evolved
517 populations at generation 103 were reared for two generations with controlled egg density (400
518 eggs/bottle) at the same temperature regime as in the evolution experiment. Freshly eclosed
519 flies were transferred onto new food for mating. Sexes were separated under CO_2 anesthesia at
520 day 3 after eclosure, left to recover from CO_2 for two days, and at the age of five days whole-
521 body mated flies of each sex were snap-frozen at 2pm in liquid nitrogen and stored at -80°C
522 until RNA extraction. More than 30 individual male flies from two reconstituted ancestral
523 populations (replicate no. 27 and no. 28) and two evolved populations (replicate no. 4 and no.
524 9) were subjected to RNA-Seq. The protocols of RNA extraction and library preparation are
525 described in (Lai and Schlötterer 2022).

526

527 RNA-Seq data analysis for mean and variance evolution

528 The processed RNA-Seq data were obtained from Lai and Schlötterer (2022). To characterize
529 the adaptive architecture of gene expression evolution, we compared the evolutionary
530 dynamics of the variance between the genes with significant mean change and those without.
531 The underlying assumption is that genes with significant mean expression changes are under
532 selection and the rest of the transcriptome is not responding to the new environment (neutral).
533 The genes with/without significant mean evolution for two evolved populations were taken
534 from Lai and Schlötterer (2022).

535 We quantified the variance changes during adaptation for each gene by using the variance
536 estimates of the expression (logCPM) of each gene in each population from Lai and Schlötterer
537 (2022). Briefly, raw read counts of each gene were normalized with the TMM method
538 implemented in edgeR. We then applied natural log transformation to the expression of each
539 gene (counts per million (CPM)) to fit normal assumption for all subsequent analyses and make
540 mean and variance independent from each other.

541 Due to the moderate sample size, we performed additional checks for the uncertainty of
542 variance estimates. Jackknifing was applied to measure the uncertainty of estimator (Fukunaga
543 & Hummels, 1989). The procedure was conducted independently on four populations, and we
544 calculated the 95% confidence interval of the estimated variance (Supplementary Figure S2).
545 Additionally, the robustness of variance estimation was also supported by the high correlation
546 (rho = 0.79, Spearman's rank correlation) in variance estimates between the two independently

547 reconstituted ancestral populations. The change of gene expression variance was determined
548 by the F statistics calculated as the ratio between the variance within the ancestral population
549 and the variance within the evolved population of each gene. To test whether selection on mean
550 expression generally alters the expression variance, we compared the F statistics of genes with
551 significant changes in mean expression to the genes without.

552

553 We note that the natural log transformation (as in Lai and Schlötterer 2022) does not attempt
554 to quantify and partition the noise due to the read sampling process at the gene level from the
555 across-individual variance. Since we were primarily interested in the latter, we evaluated the
556 effect of read sampling noise. We did an additional check by distinguishing the true biological
557 variation and the measurement error ([eq1], from User's Guide of edgeR) of each gene using
558 the statistical method implemented in edgeR, where the true variance across individuals
559 (biological coefficient of variation (BCV)) of each gene can be estimated (tag-wised
560 dispersion).

561 $$\mathrm{T}otal\ CV^2 = Technical\ CV^2 + Biological\ CV^2\ [eq1]$$

562 The proportion of technical CV^2 in total CV^2 was calculated. We showed that the average
563 proportion of technical CV^2 gradually decreased with increasing sample size (Supplementary
564 Figure S3), and for our sample size (n = 22) it is 17%. The correlation between BCV^2 and the
565 variance estimates after log transformation is 0.98 (Spearman's rho), suggesting that the two
566 estimates are similar. As a sanity check, we repeated the comparisons of the variance changes
567 in selected (DE) and neutral (non-DE) genes using BCV^2. Similarly, the changes in variance
568 of putative adaptive genes are indistinguishable from the genes that do not change their mean
569 expression (t-test, p-value > 0.05; Supplementary Figure S3).

570

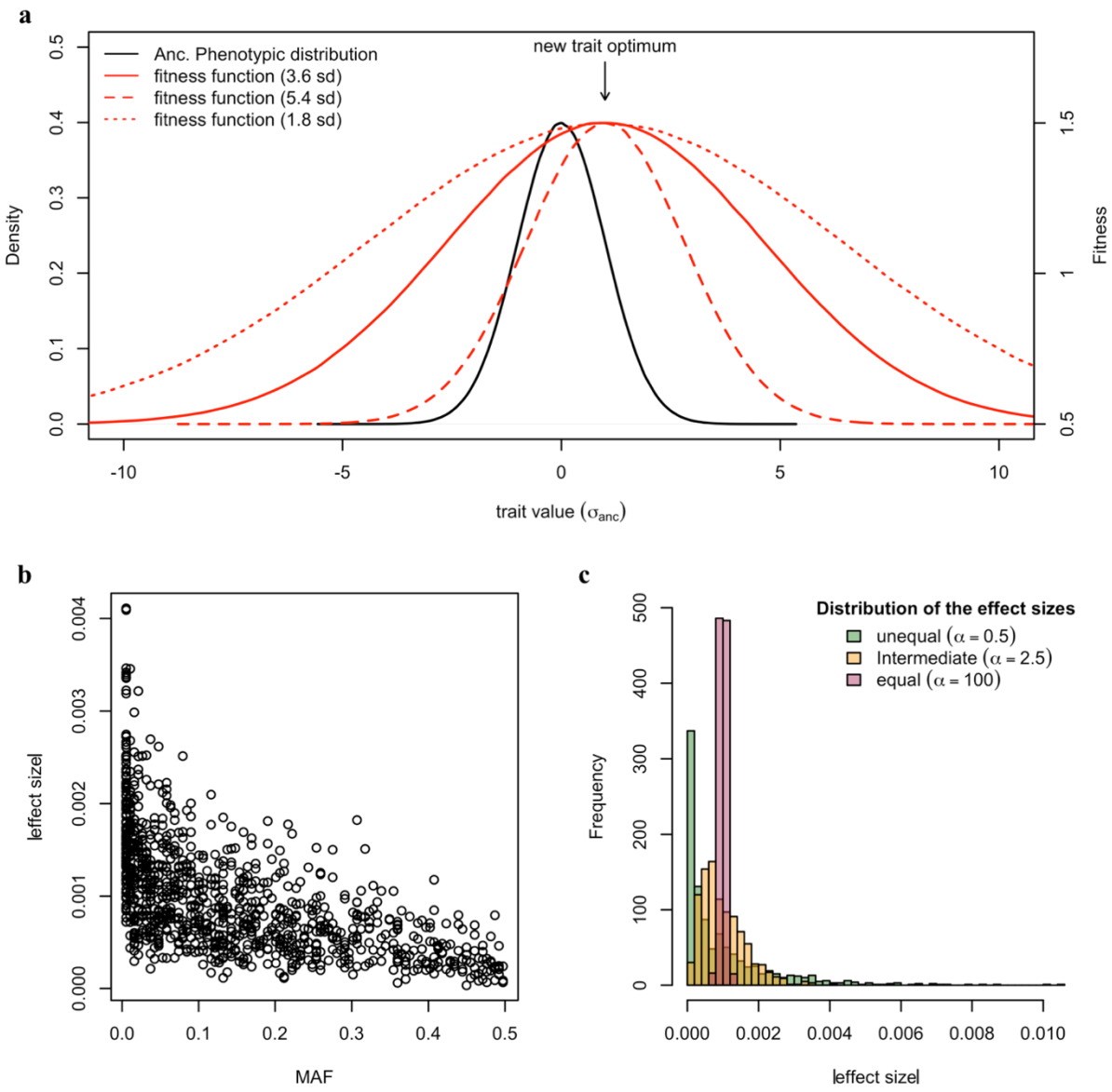


Figure 1. Simulating polygenic adaptation to a shift in trait optimum with different parameter combinations. a. For the computer simulations we consider a quantitative trait (in black) experiencing a sudden shift in trait optimum under stabilizing selection. The underlying fitness functions are illustrated 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 Otte *et al.* (2020)). We consider such negative correlation when assigning the effect sizes to variants underlying a simulated trait. **c.** The distribution of effect sizes of the contributing loci is determined by the shape parameter (α) of gamma sampling process (α = 0.5, 2.5 and 100).

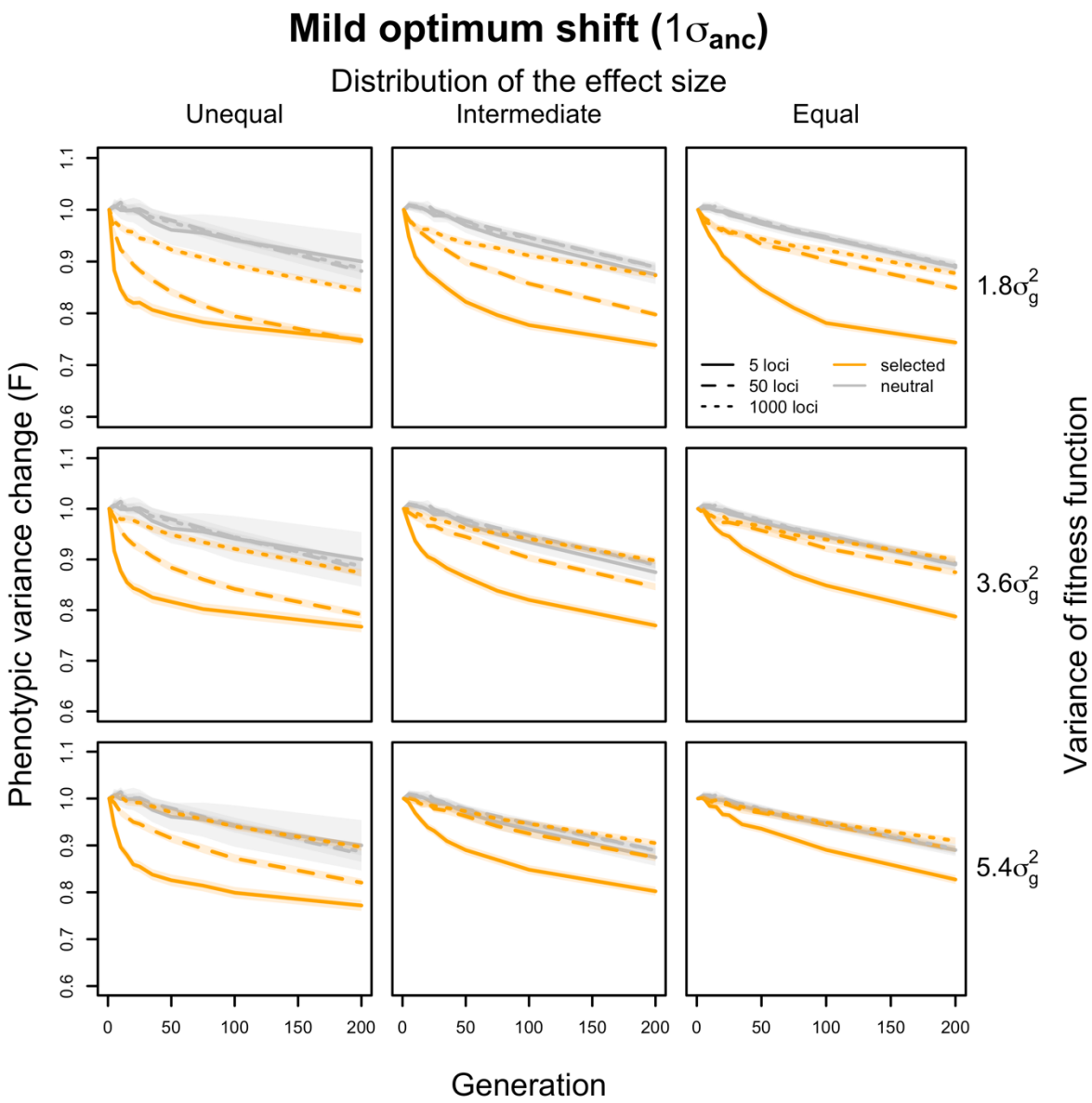
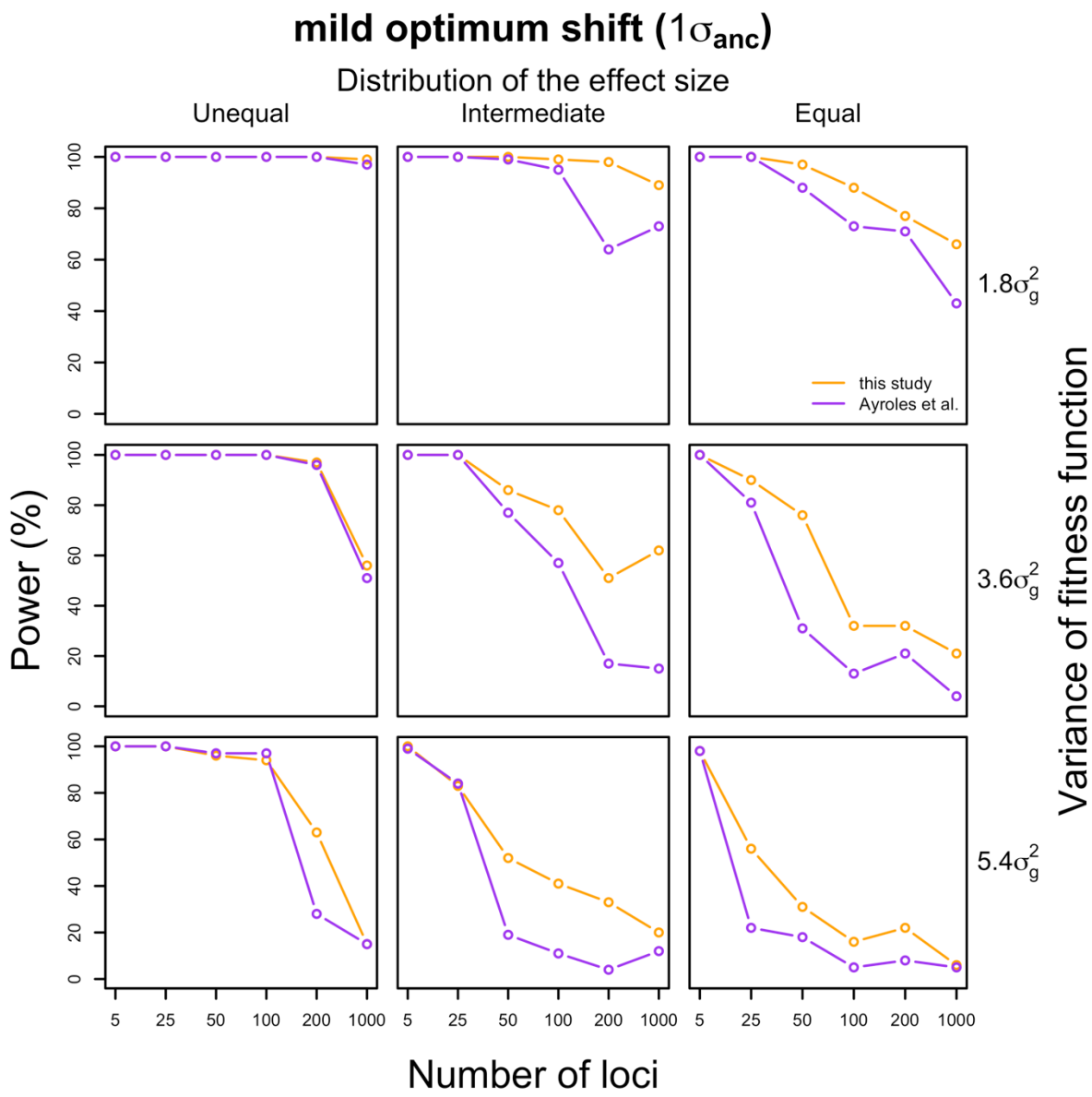


Figure 2. The trajectory of changes in phenotypic variance during adaptation to a mild optimum shift. The changes in phenotypic variance within 200 generations adapting to a moderate optimum shift (orange) are compared to the changes under neutrality (grey) on y axis. The average variance changes (F) of 1000 simulated traits is calculated as the ratio of phenotypic variance between each evolved time point (generation x) and the ancestral state (σ_x^2/σ_1^2). The translucent band indicates the 95% confidence interval for 1000 simulated traits. 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). 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

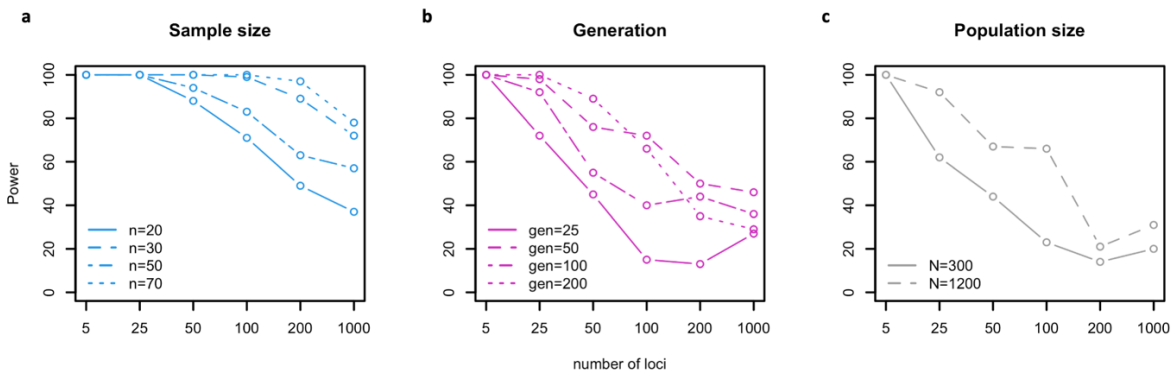
number of loci (orange solid lines; 5 loci). While, for traits with extremely polygenic basis, the phenotypic variance stays stable over time (orange dotted lines).



598

599 **Figure 3. Power for detecting significant variance change between groups of selected and**
600 **neutral traits.** For each set of the 1000 traits controlled by different numbers of loci (x-axis)
601 with varying effect sizes (columns) under each selection strength (rows), we calculated how
602 often no difference in variance change is detected between 1000 neutral and 1000 selected
603 traits after 100 generations (y-axis). The same genetic architecture is assumed for all 1000
604 selected traits with a sample size of 20. For each parameter combination, we have almost 100%
605 power when all traits were controlled by few loci (5 loci). It gradually increases with an
606 increasing number of loci (red). A lack of significant difference indicates polygenic
607 architecture of the adaptive traits. The results remain unaffected when we used heritability
608 estimates from Ayroles et al., (2009) (purple).

609



610

611 **Figure 4. Power for detecting significant variance change between groups of selected and**
612 **neutral traits under different a. sample sizes, b. generations and c. population sizes.** For
613 each set of the 1000 traits controlled by different numbers of loci (x-axis) under different
614 experimental conditions, we calculated how often a difference in variance change is detected
615 between 1000 neutral and 1000 selected traits (y-axis). The same genetic architecture is
616 assumed for all 1000 selected traits. **a.** The power gradually increases with larger sample sizes.
617 **b.** Given the same number of loci; a higher power is obtained when the experiment continues
618 for more generations. **c.** Experiments with larger population size (N=1200) have a higher power
619 than those with a smaller population size (N=300). Overall, larger and longer evolution
620 experiments with more phenotyped samples provide more power to reject simpler adaptive
621 architectures and thus provide more confidence in a highly polygenic architecture when we
622 observed no significant difference in variance.

623

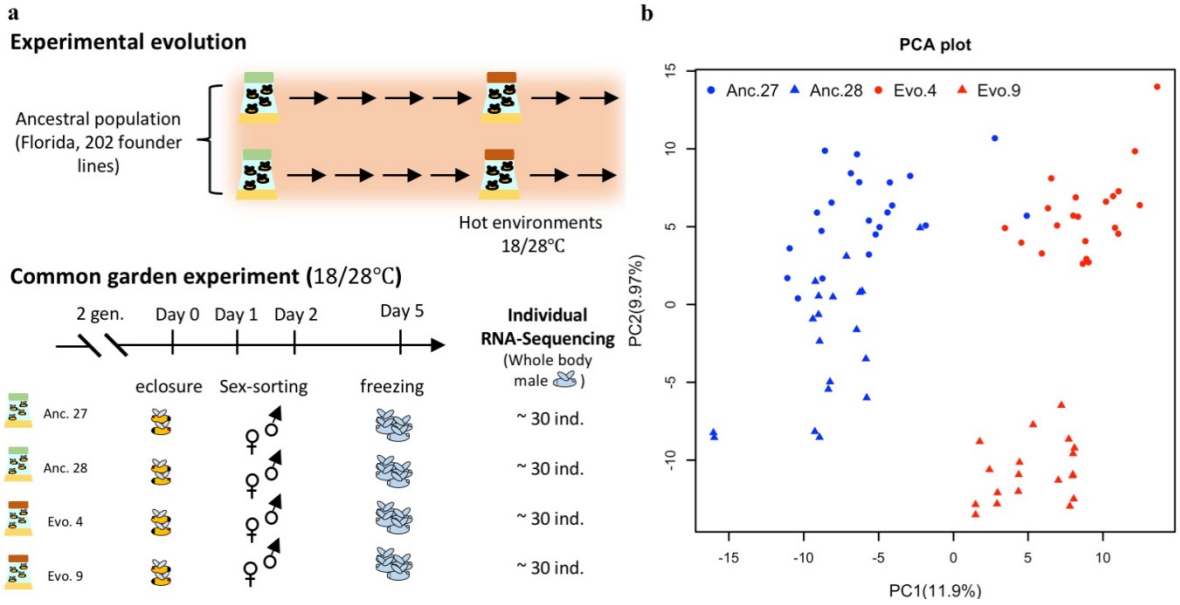
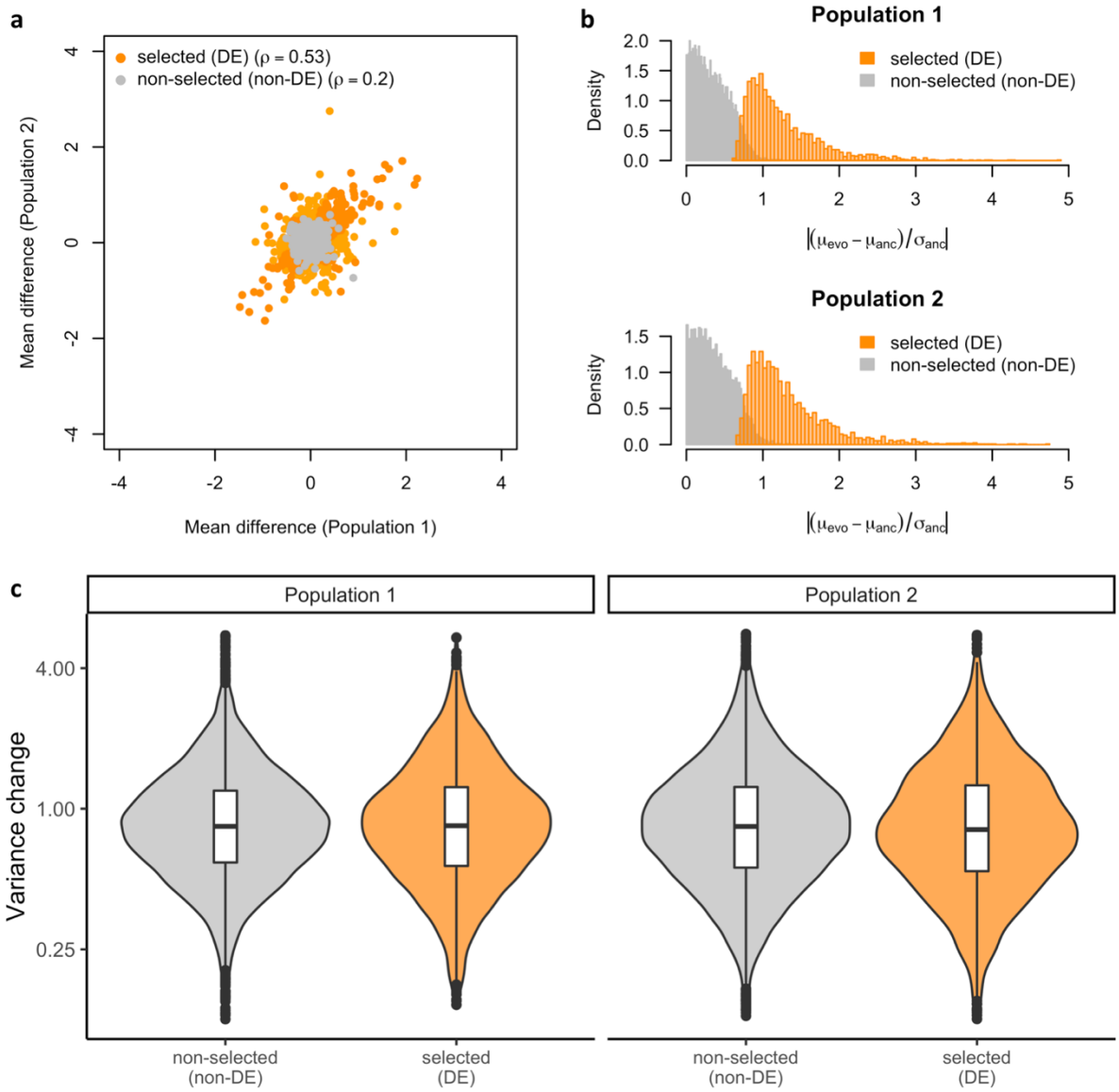


Figure 5. Schematic overview of the experimental procedures (a) and the divergence in gene expression during experimental evolution (b). a. Experimental evolution: starting from one common founder population, two replicate populations evolved for 100 generations in a hot laboratory environment fluctuating 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, males from each population were subjected to RNA-Seq. **b.** Principal 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. The impact of the library card batch can be seen from the separation of the ancestral replicates which were reconstituted from the same iso-female lines.

640



641

642 **Figure 6. Evolution of phenotypic mean and variance over 100 generations of adaptation**
643 **in empirical data. a.** The evolution of gene expression means during adaptation in the two
644 replicates. For the genes with significant changes (DE, in orange), the changes are correlated
645 between replicates (Spearman's rho = 0.53). For the genes without significant changes (non-
646 DE, in grey), the correlation between replicates is much lower (Spearman's rho = 0.2). **b.** The
647 evolution of gene expression means scaled by the ancestral variation. For the DE genes (in
648 orange), the median change is around one standard deviation of the ancestral expression value,
649 suggesting mild shift in trait optimum in the novel environment. For the non-DE genes (in
650 grey), the changes in expression are mostly negligible. The same pattern is seen in the second
651 replicate. **c.** The change in expression variance during adaptation for DE and non-DE genes. In

both replicates, the distribution of variance changes is indistinguishable between DE genes (orange) and non-DE genes (grey) (t-test, $p > 0.05$ for both replicates). The almost identical pattern is observed when the biological coefficient of variation (BCV; see materials and methods) is used to estimate the expression variance (Supplementary Figure S4).

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, David Houle, Dagný Ásta Rúnarsdóttir and anonymous reviewers 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.

Data accessibility statement

All sequencing data will be available in European Nucleotide Archive (ENA) under the accession number PRJEB37011 upon publication.

Competing interests

The authors declare no competing interests.

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

References

Ayroles, J.F., Carbone, M.A., Stone, E.A., Jordan, K.W., Lyman, R.F., Magwire, M.M., *et al.* (2009) Systems genetics of complex traits in Drosophila melanogaster. *Nature Genetics*, **41**, 299–307.

Barghi, N., Hermisson, J. & Schlötterer, C. (2020) Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics*, **21**, 769–781.

Barghi, N., Tobler, R., Nolte, V., Jakšić, A.M., Mallard, F., Otte, K.A., *et al.* (2019) Genetic redundancy fuels polygenic adaptation in Drosophila. *PLOS Biology*, **17**, e3000128.

Barton, N.H., Etheridge, A.M. & Véber, A. (2017) The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, **118**, 50–73.

Barton, N.H. & Keightley, P.D. (2002) Understanding quantitative genetic variation. *Nature Reviews Genetics*, **3**, 11–21.

Barton, N.H. & Turelli, M. (1987) Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetical Research*, **49**, 157–173.

Blair, G., Cooper, J., Coppock, A., Humphreys, M., Rudkin, A. & Fultz, N. (2019) fabricatr: Imagine Your Data Before You Collect It.

Boyle, E.A., Li, Y.I. & Pritchard, J.K. (2017) An expanded view of complex traits: From polygenic to omnigenic. *Cell*, **169**, 1177–1186.

Bulmer, M.G. (1972) The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research*, **19**, 17–25.

Burke, M.K., Dunham, J.P., Shahrestani, P., Thornton, K.R., Rose, M.R. & Long, A.D. (2010) Genome-wide analysis of a long-term evolution experiment with Drosophila. *Nature*, **467**, 587–590.

Careau, V., Wolak, M.E., Carter, P.A., Garland, T. & Jr. (2015) Evolution of the additive genetic variance-covariance matrix under continuous directional selection on a complex behavioural phenotype. *Proceedings. Biological sciences*, **282**.

Castle, W.E. (1921) An improved method of estimating the number of genetic factors concerned in the cases of blending inheritance. *Science (New York, N.Y.)*, **54**, 223.

Charlesworth, B. & Charlesworth, D. (2010) *Elements of Evolutionary Genetics*. Greenwood Village, CO: Roberts and Company Publishers.

Chevalet, C. (1994) *An approximate theory of selection assuming a finite number of quantitative trait loci*. *Genetics Selection Evolution*.

Chippindale, A.K., Chu, T.J.F. & Rose, M.R. (1996) Complex trade-offs and the evolution of starvation resistance in Drosophila melanogaster. *Evolution*, **50**, 753–766.

Clutton-Brock, T.H. & Pemberton, J.M. (Josephine M. (2004) *Soay sheep : population dynamics and selection on St. Kilda*. Cambridge University Press.

Daborn, P.J., Yen, J.L., Bogwitz, M.R., Goff, G. Le, Feil, E., Jeffers, S., *et al.* (2002) A single p450 allele associated with insecticide resistance in Drosophila. *Science (New York, N.Y.)*, **297**, 2253–6.

Endler, L., Gibert, J.-M., Nolte, V. & Schlötterer, C. (2018) Pleiotropic effects of regulatory variation in *tan* result in correlation of two pigmentation traits in *Drosophila melanogaster*. *Molecular Ecology*, **27**, 3207–3218.

Falconer, D.S. & Mackay, T.F.C. (1963) *Introduction to quantitative genetics*. *Poultry Science*. CreateSpace Independent Publishing Platform.

Ferea, T.L., Botstein, D., Brown, P.O. & Rosenzweig, R.F. (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 9721–6.

Franssen, S.U., Kofler, R. & Schlötterer, C. (2017) Uncovering the genetic signature of quantitative trait evolution with replicated time series data. *Heredity*, **118**, 42–51.

Fukunaga, K. & Hummels, D.M. (1989) Leave-one-out procedures for nonparametric error estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 421–423.

Hayward, L.K. & Sella, G. (2019) Polygenic adaptation after a sudden change in environment. *bioRxiv*, 792952.

Hill, M.S., Zande, P. Vande & Wittkopp, P.J. (2021) Molecular and evolutionary processes generating variation in gene expression. *Nature Reviews Genetics*, **22**, 203–215.

Howie, J.M., Mazzucco, R., Taus, T., Nolte, V. & Schlötterer, C. (2019) DNA motifs are not general predictors of recombination in two drosophila sister species. *Genome Biology and*

Evolution, **11**, 1345–1357.

Hsu, S.-K., Jakšić, A.M., Nolte, V., Barghi, N., Mallard, F., Otte, K.A., *et al.* (2019) A 24 h age difference causes twice as much gene expression divergence as 100 generations of adaptation to a novel environment. *Genes*, **10**, 89.

Hsu, S.-K., Jakšić, A.M., Nolte, V., Lirakis, M., Kofler, R., Barghi, N., *et al.* (2020) Rapid sex-specific adaptation to high temperature in drosophila. *eLife*, **9**.

Huang, Y. & Agrawal, A.F. (2016) Experimental evolution of gene expression and plasticity in alternative selective regimes. *PLOS Genetics*, **12**, e1006336.

Jain, K. & Stephan, W. (2015) Response of polygenic traits under stabilizing selection and mutation when loci have unequal effects. *G3 (Bethesda, Md.)*, **5**, 1065–74.

Jakšić, A.M., Karner, J., Nolte, V., Hsu, S.-K., Barghi, N., Mallard, F., *et al.* (2020) Neuronal function and dopamine signaling evolve at high temperature in Drosophila. *Molecular Biology and Evolution*, **37**, 2630–2640.

Kawecki, T.J., Lenski, R.E., Ebert, D., Hollis, B., Olivieri, I. & Whitlock, M.C. (2012) Experimental evolution. *Trends in Ecology & Evolution*, **27**, 547–560.

Keightley, P.D. & Hill, W.G. (1989) Quantitative genetic variability maintained by mutation-stabilizing selection balance: sampling variation and response to subsequent directional selection. *Genetical Research*, **54**, 45–58.

Kimura, M. & Crow, J.F. (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725–738.

Lai, W.-Y. & Schlötterer, C. (2021) Evolution of gene expression variance during adaptation to high temperature in Drosophila. *bioRxiv*, 2021.01.19.427270.

Lai, W. & Schlötterer, C. (2022) Evolution of phenotypic variance in response to a novel hot environment. *Molecular Ecology*, **31**, 934–945.

Lande, R. (1976) Natural selection and random genetic drift in phenotypic evoltion. *Evolution*, **30**, 314–334.

Lenski, R.E., Travisano, M., Larison, B. & Moritz, C. (1994) Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences*, **91**, 6808–6814.

767 Liu, X., Li, Y.I. & Pritchard, J.K. (2019) Trans effects on gene expression can drive
768 omnigenic inheritance. *Cell*, **177**, 1022-1034.e6.

769 Mackay, T.F.C., Stone, E.A. & Ayroles, J.F. (2009) The genetics of quantitative traits:
770 challenges and prospects. *Nature Reviews Genetics*, **10**, 565–577.

771 Mallard, F., Nolte, V., Tobler, R., Kapun, M. & Schlötterer, C. (2018) A simple genetic basis
772 of adaptation to a novel thermal environment results in complex metabolic rewiring in
773 Drosophila. *Genome Biology*, **19**, 119.

774 Nielsen, R. (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**,
775 197–218.

776 Nouhaud, P., Tobler, R., Nolte, V. & Schlötterer, C. (2016) Ancestral population
777 reconstitution from isofemale lines as a tool for experimental evolution. *Ecology and*
778 *Evolution*, **6**, 7169–7175.

779 Otte, K.A., Nolte, V., Mallard, F. & Schlötterer, C. (2020) The adaptive architecture is
780 shaped by population ancestry and not by selection regime. *bioRxiv*, 2020.06.25.170878.

781 Pélabon, C., Hansen, T.F., Carter, A.J.R. & Houle, D. (2010) Evolution of variation and
782 variability under fluctuating, stabilizing, and disruptive selection. *Evolution*, **64**, 1912–25.

783 Penna, A., Melo, D., Bernardi, S., Oyarzabal, M.I. & Marroig, G. (2017) The evolution of
784 phenotypic integration: How directional selection reshapes covariation in mice. *Evolution*,
785 **71**, 2370–2380.

786 Pritchard, J.K., Pickrell, J.K. & Coop, G. (2010) The genetics of human adaptation: Hard
787 sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208–R215.

788 Romero, I.G., Ruvinsky, I. & Gilad, Y. (2012) Comparative studies of gene expression and
789 the evolution of gene regulation. *Nature Reviews Genetics*, **13**, 505–516.

790 Signor, S.A. & Nuzhdin, S. V. (2018) The evolution of gene expression in cis and trans.
791 *Trends in Genetics*, **34**, 532–544.

792 Sork, V.L. (2017) Genomic studies of local adaptation in natural plant populations. *Journal*
793 *of Heredity*, **109**, 3–15.

794 Suvorov, A., Nolte, V., Pandey, R.V., Franssen, S.U., Futschik, A. & Schlötterer, C. (2013)
795 Intra-specific regulatory variation in Drosophila pseudoobscura. *PLoS ONE*, **8**, e83547.

Turelli, M. (1984) Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology*, **25**, 138–193.

Vlachos, C. & Kofler, R. (2018) MimicrEE2: Genome-wide forward simulations of Evolve and Resequencing studies. *PLOS Computational Biology*, **14**, e1006413.

Walsh, B. & Lynch, M. (2018) Evolution and Selection of Quantitative Traits, **1**.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. (2004) Evolutionary changes in cis and trans gene regulation. *Nature*, **430**, 85–88.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. (2008) Regulatory changes underlying expression differences within and between Drosophila species. *Nature Genetics*, **40**, 346–350.

Yoo, B.H., Nicholas, F.W. & Rathie, K.A. (1980) Long-term selection for a quantitative character in large replicate populations of Drosophila melanogaster. *Theoretical and Applied Genetics*, **57**, 113–117.