Gene expression evolution during adaptation is predicted by stronger genetic selection at more pleiotropic genes in *Tribolium castaneum*.

Eva L. Koch[1,2]*, Charles Rocabert[3,4]*, Champak Beeravolu Reddy[1,5], Frédéric Guillaume[1,3,†]

[1] Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland

[2] Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom

[3] Organismal and Evolutionary Biology Research Programme, University of Helsinki, 00014 Helsinki, Finland

[4] Institute for Computer Science and Department of Biology, Heinrich-Heine Universität, 40221 Düsseldorf, Germany

[5] Scientific IT, Université de Fribourg, 1700 Fribourg, Switzerland

[*] E.L.K. and C.R. contributed equally to this work

[†] Corresponding author:
Frédéric Guillaume
University of Helsinki
Viikinkaari 1
P.O. Box 65
00014 Helsinki
email: frederic.guillaume@helsinki.fi
phone: +358 50 409 80 61

keywords: experimental evolution, evolvability, G-matrix, Robertson-Price-equation, gene co-expression network, RNA-seq

Abstract

Changes in gene expression play a fundamental role in the process of adaptation and can inform on the genetic basis of adaptation. We utilized transcriptome-wide variation in gene expression as a means to uncover genes under selection for expression changes during adaptation to heat and drought stress and to understand the nature of selection on gene expression traits of the red flour beetle *Tribolium castaneum*. We showed that estimates of genetic selection on transcript abundance were predictive of evolutionary changes in gene expression after 20 generations of adaptation in seven independent experimental lines. Having measured the genetic covariance between gene expression and relative fitness and among expression traits, we showed that evolutionary changes were caused by indirect selection acting on genetically correlated partners rather than direct selection acting on isolated genes. Consequently, pleiotropic genes with central positions in gene co-expression networks experienced stronger selection and exhibited larger evolutionary changes in expression. Our genomic analysis revealed that selection on expression levels drives parallel allele frequency changes (AFC) in the respective genes. More pleiotropic genes and those carrying expression quantitative trait loci (eQTLs) showed a higher degree of parallel evolution. More generally, genes with more parallel AFCs were under stronger genetic selection. Contrary to previous evidence of constrained evolution at more connected genes, adaptation was driven by selection acting disproportionately on genes central to co-expression gene networks. We demonstrated that measures of selection at the transcriptome level can provide accurate evolutionary predictions and critical information on the molecular basis of rapid adaptation.

Introduction

Understanding the genetic basis of adaptation is a central goal in evolutionary biology. While significant progress has been made in mapping molecular variation onto phenotypic variation, we still face challenges in understanding how molecular variation translates into fitness variation to enable evolutionary changes. Molecular studies of the genetic basis of trait variation have revealed the key role played by gene expression in phenotypic variation

and unraveled networks of locally and distantly acting gene regulation affecting variation in expression (Brem et al., 2002; Wray, 2007; Stern and Orgogozo, 2008). Yet, although the role of gene expression in adaptation is largely accepted (Fay and Wittkopp, 2008) and has been amply documented (*e.g.*, Whitehead and Crawford, 2006; Fraser et al., 2010; Pritchard et al., 2017; Mack et al., 2018; Kelly et al., 2023), only few recent studies provide direct measurements of the relationship between gene expression and fitness in insects (McGuigan et al., 2014; Koch and Guillaume, 2020a), plants (Groen et al., 2020), or vertebrates (Ahmad et al., 2021). Such measurements provide estimates of phenotypic selection gradients on whole-transcriptome data from which evolutionary responses in gene expression can be predicted (Falconer and Mackay, 1996). This approach has been utilized to find genes whose expression variation is most important for fitness variation under particular stress (*e.g.*, drought in maize, Groen et al., 2020). However, we do not yet know how expression traits may actually respond to such selection pressures. The predictability of gene expression evolution thus remains untested.

Predicting evolutionary responses from phenotypic selection gradients may be biased by the presence of confounding environmental co-variation between gene expression and fitness. Gene expression can be strongly environmentally sensitive and noisy, responding plastically to abiotic stressors (Lòpez-Maury et al., 2008; Hodgins-Davis et al., 2012; des Marais et al., 2013). Although this plasticity may allow for rapid physiological and metabolic adjustments to new environmental conditions (Gibson, 2008; Ghalambor et al., 2015; Koch and Guillaume, 2020a) and enable adaptive changes when genetically grounded (des Marais et al., 2017), it may also result in non-genetic causes of phenotypic covariation with fitness. A response to selection may then be wrongly expected when phenotypic covariation obscures genetic covariation with fitness (Kruuk et al., 2002; Morrissey et al., 2010; Bonnet et al., 2017). More accurate predictions of the selection responses of phenotypic traits, including gene expression traits, are based on estimates of genetic selection gradients provided by the application of the Robertson-Price equation (or Secondary theorem of natural selection Robertson, 1966; Price, 1970). The method requires to estimate the genetic covariation of additive genetic variation of the trait and fitness and has been advocated to study evolution in wild populations (Morrissey et al., 2010, 2012) among other applications.

One important aspect of the estimates of genetic covariation with fitness is that they include the effects of direct and indirect selection on the trait of interest, where indirect selection stems from direct selection on the traits correlated to the focal trait. The estimate of total selection provided by the Robertson-Price relation averages out the potential constraining and enhancing effects of selection on multiple traits (Walsh and Blows, 2009). Evolutionary changes might then occur because a trait is the direct target of selection, or because it is genetically correlated to traits under selection, or both. Disentangling effects of direct and indirect selection, although challenging, is especially relevant for gene expression phenotypes. It would help us gain detailed insights into mechanisms of how and why certain expression levels change, and understand their exact role in adaptation. Since genetic covariances between gene expression levels are prevalent (Ayroles et al., 2009; McGraw et al., 2011; Blows et al., 2015), the evolutionary response of a single gene's expression may thus be more influenced by indirect than direct selection.

Pleiotropy, the property of genes affecting multiple traits, plays a key role in generating genetic and phenotypic covariance among traits. The importance of genetic covariance is often discussed in the context of evolutionary constraints and evolvability (Blows and Hoffmann, 2005; Walsh and Blows, 2009; Hansen, 2003; Hansen and Pelabon, 2021). The presence of genetic covariance implies that a trait cannot evolve independently, as selection will also affect correlated traits, potentially preventing responses to certain combinations of selective pressures. However, if selection acting on multiple traits is aligned with the direction of their genetic covariance, it can enhance selection responses and accelerate evolution. In this scenario, a single trait experiences not only the direct effect of selection but also the combined effects of selection acting on all its genetically correlated partners (Lande, 1979; Agrawal and Stinchcombe, 2009).

The evolutionary importance and potential constraints associated with covariances between expression levels are largely unknown. For now, experimental evolution studies demonstrated rapid evolution of gene expression during adaptation to new environments, suggesting that standing variation in expression is relatively unconstrained (Yampolsky et al., 2012; Ghalambor et al., 2015; Huang and Agrawal, 2016; Koch and Guillaume, 2020b), and mutational accumulation experiments showed that the influx of mutational variation in gene expression can be large (Rifkin et al., 2005). However, studies of the

genetic variance-covariance structure of gene expression show that species divergence in expression is constrained by genetic interactions (Fraser et al., 2002) and correlations (Innocenti and Chenoweth, 2013) within gene networks. Gene expression variation also seems predominantly under stabilizing selection (Josephs et al., 2015; Kita et al., 2017), potentially constraining genetic variation at pleiotropic genes, but other studies suggest that the constraining effect of stabilizing selection on expression variation might be weak (Steige et al., 2017; Glassberg et al., 2019). All in all, the potential for evolution of gene expression is likely large but might remain constrained by the regulatory and functional network context of variation in expression (Jordan et al., 2004; Innocenti and Chenoweth, 2013; Collet et al., 2018). Currently, emerging evidence from experimental evolution experiments and studies of the genetic basis of local adaptation to environmental variation suggest that pleiotropy (or gene connectivity in gene networks; He and Zhang, 2006) plays a key positive rather than constraining role in adaptation (Rennison and Peichel, 2022; Ruelens et al., 2023; Thorhölludottir et al., 2023; Whiting et al., 2024).

Here, we took advantage of individual whole transcriptome (RNA-seq) and fitness data from two time points within replicated experimental evolution lines of *Tribolium castaneum* (red flour beetle) adapting to a stressful hot-dry treatment. Combining expression and fitness data (offspring number) of 486 female individuals in a paternal half-sib breeding design enabled us to estimate genetic covariances between expression levels and fitness in the first generation and to make predictions about evolutionary changes. A previous study demonstrated that adaptation involved significant changes in gene expression (Koch and Guillaume, 2020b). Here, our goal was to evaluate the predictability of gene expression evolution by assessing whether long-term (20 generations) evolutionary changes were consistent with short-term predictions (one generation). We further wanted to understand in detail how and why expression evolved and quantified the contributions of direct and indirect selection to evolutionary changes. We built on previous studies that developed methods to estimate the genetic covariance matrix of gene expression in such high dimensionality data by using a Bayesian sparse factor analysis (Runcie and Mukherjee, 2013) and disentangle the contributions of direct and indirect selection to the expected evolutionary change (Stinchcombe et al., 2014).

Additionally, constructing gene co-expression networks enabled us to explore how ex-

pression network structure changes under stress. Combining it with selection estimates gave us insights into how selection was acting on different network components and how this was then ultimately related to improving fitness under stressful conditions. Using this approach, we could address crucial questions regarding the genetic architecture of gene expression adaptation: Are there few genes that are direct targets of selection with the majority of genes showing correlated response? Is selection acting on many gene expression levels simultaneously, and can position in gene expression networks predict their response to selection?

Finally, we wanted to link evolutionary changes in expression levels with genetic changes. Having seven replicate selection lines we tested for parallel allele frequency changes (AFC) and whether these changes were associated with genes under selection for expression changes. Furthermore, we identified eQTLs, SNPs associated with variation in expression patterns, and tested whether these showed parallel AFC in the different selection lines.

Results

Evolutionary changes are consistent with predictions

Significant additive genetic variance (V_A) in fitness is the key requirement for any evolutionary change and gives the evolutionary potential for adaptation (Fisher, 1930). The additive genetic variance of mean-standardized relative fitness ($I_A(w) = V_A(w)/\bar{w}^2$; w: relative fitness, $\bar{w}$: mean relative fitness) showed significant evolutionary potential for increased fitness in the hot-dry treatment (HD) with $I_A(w) = 0.033$, higher than in the control condition (CT) ($I_A(w) = 0.019$). This is consistent with the expectation that additive genetic variance in relative fitness $V_A(w)$ should be higher in novel conditions compared to ancestral conditions where strong selection in the past has removed genetic variation. It also agrees with the previously observed adaptation in the selection lines (Koch and Guillaume, 2020b). Similarly, evolutionary changes in the mean expression level of a gene requires additive genetic variation in expression. Genes greatly varied in their proportion of additive genetic (V_A) to phenotypic variance (V_P), as shown by the distribution of their heritability coefficients ($h^2 = \frac{V_A}{V_P}$; Figure S1). Median h^2 in HD was

0.64, very close but significantly lower than in CT (0.66, $P < 0.0001$, permutation test), indicating a large capacity to respond to selection.

We evaluated the evolvability of gene expression in HD with estimates of the genetic covariance of gene expression with relative fitness: $s_g = cov_a(w, z)$, hereafter called genetic selection. The s_g coefficient is also the predicted response to selection of the trait z (here gene expression) over one generation (Robertson, 1966; Price, 1970). We confronted those predictions with the changes in gene expression observed after 20 generations of adaptation in HD to assess the predictability of our estimates of genetic selection.

We found that we were able to predict changes in gene expression based on the genetic covariances of single gene expression levels with fitness. We detected a significant correlation between the transcriptome-wide responses predicted by the genetic covariance with fitness (s_g) in HD and the observed evolutionary changes in expression ($\rho = 0.22$; bootstrap 95% CI: [0.20, 0.24]; permutation test: $P < 0.0001$). 55 gene showed a significant s_g and 75 genes showed significant r_g (genetic correlation, *i.e.*, variance-standardized s_g) after correction for multiple testing. When we focused on the set of genes with significant s_g, the correlation with observed evolutionary changes becomes even stronger ($N = 55$: $\rho = 0.64$, bootstrap 95% CI: [0.38, 0.83]; permutation test: $P < 0.0001$). Genes under positive genetic selection ($s_g > 0$) had on average significantly increased expression levels (0.067 log2-Fold change; bootstrap 95% CI: [0.036, 0.102]), and genes under negative genetic selection ($s_g < 0$) showed reduced expression (-0.094 log2-Fold change; bootstrap 95% CI [-0.133, -0.053]; Figure 1A). The predictability of the direction and, to some extant, the magnitude of the response to selection in the HD environment was thus remarkably good (Figure 1B).

Predicted response is mainly caused by indirect selection

Separation of direct and indirect selection can be done if the genetic co-variation among traits is also known (Stinchcombe et al., 2014), achieved by using a multivariate quantitative genetics approach to estimate the G-matrix, the genetic variance-covariance matrix (Lande, 1979; Lande and Arnold, 1983). Having computed estimates of G-matrices for gene expression traits, we could compare the expected response due to direct selection $R_\beta = \beta_g * diag(\mathbf{G})$ with s_g (predicted total response), where $diag(\mathbf{G})$ is the vector of addi-

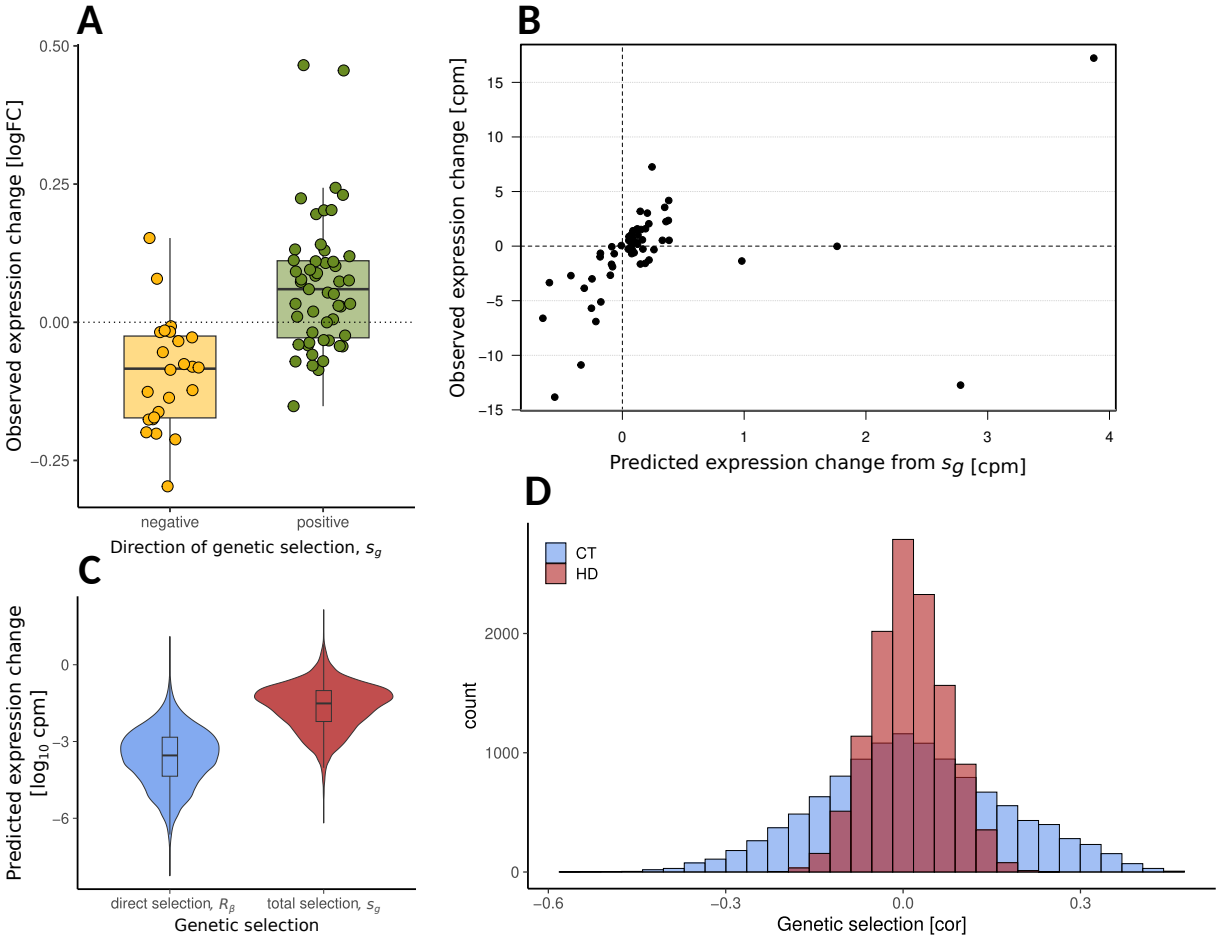


Figure 1: (**A**) Evolutionary response of gene expression ($\log_2$ fold change—logFC) in Hot-Dry conditions of the genes having significant genetic correlation with fitness in the first generation (genetic selection). The evolutionary response is the difference between selection lines after 20 generations of adaptation and non-adapted control lines in Hot-Dry conditions. (**B**) Relationship between predicted responses to genetic selection based on significant estimates of s_g and observed evolutionary changes after 20 generations of selection in Hot-Dry conditions (in counts per million gene expression levels—cpm). (**C**) Comparison of the absolute responses to direct (R_β) and total (s_g) selection in Hot-Dry conditions at the first generation for all 11,890 genes with expression data. (**D**) Distribution of the variance-standardized genetic selection coefficients (genetic correlation r_g) in the Control (CT) and Hot-Dry (HD) environments at generation 1.

tive genetic variance of gene expression. In general, we found largely weaker estimates of R_β than s_g (Figure 1C). Only 86 genes (0.7%) showed a larger magnitude of response to direct selection R_β than the predicted response s_g in HD. 25% of the genes had opposite signs of R_β than s_g. The proportion of the predicted response contributed by direct selection (R_β/s_g) in those genes where R_β and s_g were of the same sign was small with median proportion of 1%. These results indicate that indirect selection constitutes the majority of genetic selection s_g and that constraints on expression evolution due to opposing signs of direct and indirect evolution are small.

Gene co-expression networks and selection

The prevalent role of indirect selection underscores the importance of genetic correlations between expression levels and their interdependence. Analysis of gene co-expression networks is a powerful way to assess the importance of correlations among gene responses and identify clusters of co-regulated genes. A weighted gene co-expression network analysis (WGCNA; Langfelder and Horvath 2008) on gene expression levels in HD uncovered 26 modules of co-expressed genes (Figure 2A) and nine co-expression modules in CT. Overall, CT modules were preserved in HD, *i.e.*, there were no fundamental differences in the way gene expression levels are correlated with each other (Figure 2D). However, it seems that correlations between expression levels became weaker in HD leading to a splitting of the CT modules. We tested each module's fitness association by correlating the module's eigengene (corresponding to the first principal component of each module) with fitness. Six CT modules were significantly associated with fitness variation, but none of the HD modules were (Figure 2C). Additionally, when comparing selection coefficients of single gene expression levels between CT and HD, we observed a decrease in the number of significant coefficients from 729 in CT to 55 in HD, and a narrower distribution in HD relative to CT (Figure 1D), pointing to weaker genetic selection in HD. Therefore, it seems that the stress imposed by the HD conditions caused a weakening of expression-fitness and module-fitness associations, alongside a decrease of mean fitness in HD (from 49.1 in CT to 25.7 surviving offspring per female, see Koch and Guillaume, 2020a).

Reassessing the structure of gene co-expression networks after 20 generations of adaptation, we found that the number of modules in HD was drastically reduced to two modules

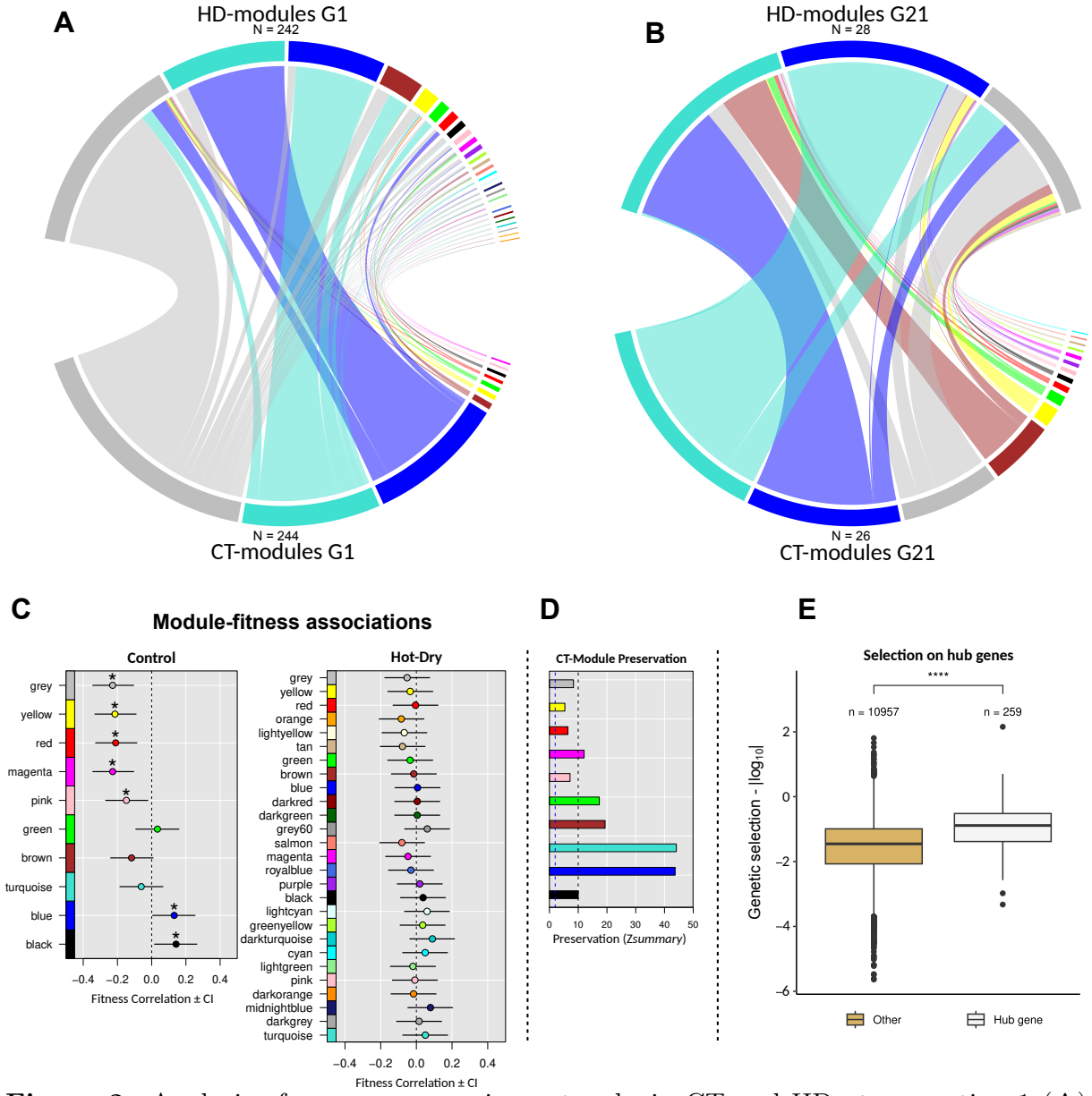


Figure 2: Analysis of gene co-expression networks in CT and HD at generation 1 (**A**) and generation 21. (**B**) The number of sampled sequences for each analysis is shown on the graphs. The significance and number of genes of module overlap are presented in Figures S3 and S4. The gray modules correspond to genes non-assigned to any cluster of co-expressed genes. (**C**) Fitness association of the gene co-expression networks in HD and CT, and (**D**) preservation of CT modules in HD at generation 1. (**E**) Difference in absolute $\log(s_g)$ of the hub versus non-hub genes.

in the adapted lines while the number of CT modules in the CT lines remained high at 14 (Figure 2B). We thus found the reverse pattern of integration of sub-networks into two large modules at generation 21 in HD. Moreover, the previous network fragmentation caused by exposing CT lines to HD conditions was somewhat recovered at generation 21 (module number increased from 14 to 19 for CT lines transplanted to HD; see Figure S6). Therefore, we can conclude that evolutionary changes which occurred during adaptation to HD conditions have led to a rewiring of gene co-expression networks.

Network position is associated with selection and evolutionary change

Despite a weaker module association with fitness at the gene network level, we found that the 260 hub genes in HD (10 most connected genes within each module, see Material and Methods) were associated with significantly higher absolute genetic selection s_g ($P < 0.0001$, permutation test, Figure 2E). Consistent with this observation, hub genes showed more evolutionary changes in expression compared to non-hub genes; they were over-represented among genes with significant evolutionary changes ($P = 1.1 \times 10^{-3}$, Fisher's exact test) and had a larger absolute evolutionary change (logFC; $P = 10^{-4}$, permutation test) in generation 21. However, direct selection on hub genes in the first generation was significantly weaker compared to other genes ($P = 10^{-4}$, permutation test).

Selection on gene expression drives parallel allele frequency changes

To better understand the importance of our estimates of selection on expression levels in driving evolutionary changes, we looked for allele frequency changes (AFCs) in the coding sequences of the expressed genes between the first and the last generation in HD, and tested whether replicated lines exhibited parallel AFCs at the same genomic positions. Parallelism of AFCs among replicated lines should provide an indication of the role of selection in enabling evolutionary changes. Consistent AFCs across replicates have been previously interpreted as evidence for the role of selection, while lack thereof would be caused by genetic drift or highly redundant genetic architecture of quantitative traits (Yeaman et al., 2018; Barghi et al., 2019, 2020). Having both AFCs across multiple experimental lines and estimates of selection on gene expression levels, we can test for a relationship between selection on gene sequence variation and gene expression.

To uncover DNA sequence variation (single nucleotide polymorphisms, SNP), we developed bio-informatics pipelines (see Figure S7) to analyze 822 whole transcriptome RNA-seq samples. After quality check, we found 356,193 SNPs in 11,216 genes suitable for the analysis (31.76 SNP per gene on average; Figure S8), uniformly distributed on the genome (Figure S9), and 958 eQTLs (located in 564 genes) significantly associated with the expression level of 556 distinct other genes. All the eQTLs were in *trans*, *i.e.*, no eQTL was locally associated to its carrier gene expression level.

Genome-wide, 55,063 SNPs showed significant AFC in HD, when tested against the AFC distribution in CT (see Methods). We then defined the parallelism of a gene as the largest parallelism of its SNPs. This resulted in 8,906 genes being classified as exhibiting AFC in at least one line. A small set of 179 SNPs shifted in at least five lines and was considered highly parallel. This subset of SNPs belonged to 147 genes. Only one SNP shifted in all seven lines sampled in HD (see Figure S10A). The correlation between $log(s_g)$ and gene parallelism was positive and significant ($\rho = 0.215$, *adj.* $R^2 = 0.046$, p-value $< 2.2e\text{-}16$). Genes binned into different categories of AFC ("no shift"; "single shift"; "partially or fully parallel shifts", Figure 3A) significantly differed in their mean s_g (all two-sided Mann-Whitney-U tests $P < 1 \times 10^{-4}$), and genes with significant s_g were significantly enriched in highly parallel SNPs (AFC in ≥ 5 lines; $P = 5 \times 10^{-3}$, permutation test on 2,245 SNPs in 75 genes minus five genes not carrying any SNP). These results suggest that selection on gene expression levels may play a key role in driving parallelism of AFCs among replicates.

On the other hand, we did not find enrichment of parallel AFCs in the genes with either largest or significant evolutionary changes in gene expression; we found no enrichment in parallelism among 2,424 SNPs located in the first 100 genes with largest evolutionary response ($P = 0.71$, permutation test), nor among 2,250 SNPs belonging to the 53 genes with a significant evolutionary response (from the 55 genes found in Koch and Guillaume 2020b and carrying at least one SNP; $P = 0.68$, permutation test). This suggests that although selection and parallelism of AFCs appear linked, the SNPs with parallel AFC did not seem to affect the evolution of expression of the genes to which they belong.

Stronger signals of selection in hub genes and more pleiotropic eQTL carrier genes

Since in our case most selection seems indirect, we expect to find stronger signals of selection within more connected genes, *i.e.*, within hub genes or genes more pleiotropically connected to other genes. Hub genes and genes carrying eQTLs (*i.e.*, SNPs significantly associated to expression variation at other genes) are obvious candidates to also test for the effects of indirect selection on gene-level AFC parallelism. We found that the 259 hub genes in HD (ten per module; minus one gene not carrying any SNP) were significantly more parallel than the rest of the genome ($P = 1.9 \times 10^{-4}$, permutation test). Indeed, among these 259 hub genes, genes with partial (AFC in two to four lines) or full parallelism (AFC in ≥ 5 lines) were over-represented compared to genes with no significant shift (χ-squared = 21.34, df = 2, p-value = 2.323e-05). Additionally, we found that hub genes were more likely to carry an eQTL than expected by chance (χ-squared = 30.688, df = 1, p-value = 3.031e-08). This strengthens the vision that hub genes may be functionally linked with other expressed genes.

We then focused on the 564 genes carrying at least one eQTL. We found that eQTL carrier genes were more highly parallel (in ≥ 5 lines; $\chi^2 = 11.974$, df = 1, p-value = 5.396e-04). Defining the pleiotropy of one eQTL as the number of distinct phenotypes (*i.e.*, gene expression levels) associated to it, we then asked whether more pleiotropic eQTL carrier genes were under stronger genetic selection. Using three pleiotropic categories, "non-eQTL carrier", "non-pleiotropic eQTL carrier" (*i.e.*, a single phenotype association) and "pleiotropic eQTL carrier" (≥ 2, max = 41; see Figure S11), we first found that expression of eQTL carrier genes was under significantly stronger genetic selection compared to non-eQTL carriers (Figure 3B; all two-sided Mann-Whitney-U tests $P < 1 \times 10^{-2}$). Second, genetic selection s_g was significantly higher among pleiotropic than non-pleiotropic eQTL carrier genes (Figure 3B), as expected from our previous finding on the predominant role of indirect selection (Figure 1C).

Functional enrichment of eQTL carrier and hub genes

We performed a Gene Ontology (GO) category enrichment analysis to gain insights into the molecular functions of hub and eQTL carrier genes, with a 5% FDR. Apart from

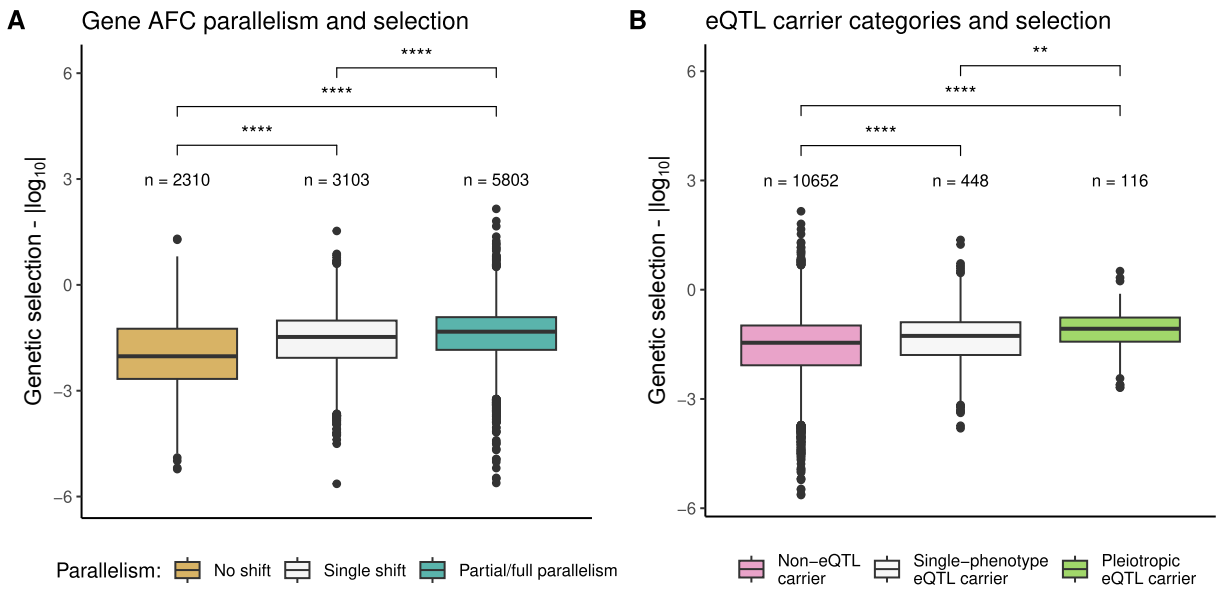


Figure 3: Genome-wide, gene AFC parallelism and eQTL carrier pleiotropy are significantly associated to selection at the gene expression level. For each panel, the distribution of genetic selection coefficients s_g (*i.e.*, covariances with fitness) is plotted by category in absolute log-scale, with all sample sizes and pairwise Mann-Whitney-U tests being shown (****: $P < 0.0001$, **: $P < 0.01$). (**A**) Gene parallelism categories: "No shift", "Single shift" and "Partial/Full parallelism". (**B**) eQTL carrier genes categories: "non-eQTL carrier", "non-pleiotropic eQTL carrier" (*i.e.*, a single phenotype association) and "pleiotropic eQTL carrier" (≥ 2, max = 41; see Figure S11).

broad functional categories like *cellular processes* or *metabolic processes*, eQTL carrier genes showed mostly enrichment in processes linked to *nuclear division* (GO : 0000280), *development* (GO : 0032502), *reproduction* (GO : 0032502) (*e.g.*, *gamete generation* (GO : 0007276), *regulation of reproductive processes* (GO : 2000241) or *meiotic cell cycle processes* (GO : 1903046)), and *stress response* (GO : 0006950). Enrichment in hub genes showed a similar pattern with *developmental processes* (GO : 0032502), *gamete generation* (GO:0007276) or *sexual reproduction* (GO : 0019953) being among the most significantly enriched categories. This underlines that processes in females are targeted towards increasing the number of offspring. Interestingly, both groups, hubs and eQTL carrier genes, showed one of the strongest enrichment for *epoxide hydrolase activity* (GO : 0004301), which was also the top enrichment category when we tested genes acting as both, hub and eQTL carrier genes. In insects, epoxide hydrolases are mostly known for their involvement in the metabolism of juvenile hormones (Daimon and Shinoda, 2013). Given the profound role of juvenile hormones as regulators of gene expression (Nijhout, 1994) and their involvement in development, reproduction and life-history trade-offs (Flatt et al.,

2005), we should expect them to be targets of selection and occupy central positions in gene regulatory networks.

Discussion

We showed that gene expression evolution is predictable and demonstrated how a quantitative genetics approach can provide insights into the nature of selection driving these evolutionary changes. Selection-driven evolutionary changes predominantly affected highly connected hub genes within co-expression modules, which may act as central regulators of gene expression in networks of co-expressed genes. Larger evolutionary changes contributing to adaptation of *T. castaneum* to hot-dry conditions were thus favored at more pleiotropic hub genes. Our study illustrates the importance of indirect selection and gene pleiotropy for understanding and predicting the evolution of correlated gene expression traits in co-expression networks.

Indirect selection drives evolutionary changes during adaptation

We showed that indirect selection is the predominant force driving evolutionary changes in gene expression during adaptation to the hot-dry conditions in our experiment. Our estimates of genetic selection predicted phenotypic responses in expression during adaptation. Disentangling direct and indirect effects contributing to genetic selection revealed that direct selection contributed very little to predicted changes in expression. This underscores that genes do not act or evolve in isolation. Instead, gene expression evolution can be better understood from a gene expression network perspective: Genes are evolving because they are highly interconnected and thus subject to high amounts of indirect selection.

Very few studies measured indirect selection acting on gene expression traits within species. Houle and Cheng (2021) showed how genetic correlation among gene expression traits may play an important role in the evolution of sexual dimorphism in expression in *Drosophila melanogaster* and Groen et al. (2020) predicted substantial constraints on responses to selection of expression traits in maize caused by indirect selection pressures of opposite direction. Among species, highly connected genes within networks were often

found to show signs of constrained sequence evolution during divergence according to their pattern of genetic co-variation (Fraser et al., 2002; Hahn and Kern, 2005; Innocenti and Chenoweth, 2013). Additionally, analyses of within population genetic variation reported purifying selection on highly connected genes (Josephs et al., 2017; Mähler et al., 2017) and predominantly stabilizing selection on gene expression variation (Josephs et al., 2015; Kita et al., 2017). Considering that the link between connectedness and pleiotropy is plausible (He and Zhang, 2006), these results are in line with the general expectation that genetic variation at more pleiotropic genes is more likely deleterious (Fisher, 1930; Orr, 2000; Otto, 2004), and more so in populations under stabilizing selection at mutation-selection balance on multidimensional phenotypic optima (Martin and Lenormand, 2006).

In contrast, our study shows that selection can lead to larger evolutionary changes at more connected genes. This outcome is also expected from models of adaptation in multidimensional trait space but when the population is away from its phenotypic optimum (Fisher, 1930; Orr, 2000; Martin and Lenormand, 2006). Selection in our experimental lines was measured at the first generation of exposure to the environmental stress and evolutionary changes were measured after 20 generations of adaptation. Our measurements were thus made during the early, less constrained phase of adaptation where we expect adaptive substitutions to have larger and more beneficial effects than when the population is at its optimum (Orr, 2000; Martin and Lenormand, 2006). Furthermore, larger adaptive steps in phenotypic space towards an optimum may be achieved by variants at more pleiotropic genes and thus favored early during adaptation. This explains why selection and evolutionary changes were stronger at hub genes in our experiment. It also explain in part why selection was generally more indirect than direct in our data, because affecting more pleiotropic genes during early adaptive steps.

Network structure increases evolvability

A further explanation for the observed positive link between evolvability and pleiotropy adding to the previous effect when a population is away from its optimum is the alignment of genetic co-variation with the direction of selection (Chevin et al., 2010). If standing genetic variation were aligned with the direction of multivariate selection towards a phenotypic optimum, then we would expect unconstrained changes at pleiotropic genes (Hansen

and Houle, 2008; Walsh and Blows, 2009). Alignment can be achieved if past selection has either sifted through existing genetic variation or selected for specific mutational correlation at pleiotropic genes (Chebib and Guillaume, 2022; Petit et al., 2023). Both phenomena would generate patterns of genetic covariation aligned with correlational selection, which then enhances selection responses and accelerates evolution in the direction of selection. This alignment is expected from theory showing that correlational selection can orient genetic covariance with the direction of selection leading to stronger responses to selection (Arnold et al., 2008; Chebib and Guillaume, 2021).

Since gene expression traits are often organized into functional modules, exhibiting little correlation with genes outside their respective groups (Ihmels et al., 2002; Barabási and Oltvai, 2004), they are likely to experience concerted selection pressures if selection acts on the functional properties of the modules resulting in correlational selection on patterns of gene co-expression. Correlational selection may thus be a driving force behind network structure (Petit et al., 2023). In line with that, several modules in the ancestral control conditions (CT) exhibited significant phenotypic correlations with offspring number, indicating selection acting on whole-module expression and alignment of covariation patterns of expression with the direction of selection. We should thus expect that selection in the past has shaped the genetic covariance between expression levels to be aligned with selection acting on them (Watson et al., 2014; Petit et al., 2023). However, in the hot-dry treatment (HD), the phenotype-fitness correlation dwindled, showing a potentially strong effect of the stress environment on the network structure. Our results suggest that the metabolic and physiological stress responses in the first generation of selection affected gene expression in such a way as to mis-align gene expression modules with the direction of selection. Yet, evolution proceeded from that existing variation and led to changes in gene expression and network structure, potentially re-aligning co-expression modules with new selection pressures. Our results further suggest that evolutionary changes at hub genes of past modules were instrumental for the re-organization of the gene co-expression networks during rapid adaptation to the new environment. Our results thus also support the idea that the topology of gene networks is dynamical (Luscombe et al., 2004) and should reflect the topology of direct and indirect selection pressures (Burban et al., 2021; Petit et al., 2023).

Predictability and parallelism of evolutionary changes

We could provide a direct test of the predictability of selection responses of gene expression traits over multiple generations, which previous studies could not. Our genetic selection estimates were informative of the phenotypic changes that occurred after 20 generations of adaptation. They were also informative of the parallelism of allele frequency changes (AFC parallelism) among experimental lines. Previous studies estimated either phenotypic selection on expression traits (selection differentials; McGuigan et al., 2014; Groen et al., 2020; Koch and Guillaume, 2020a; Ahmad et al., 2021) or AFC parallelism in sequence data from replicated populations (e.g., Bergland et al., 2014; Barghi et al., 2019; Rudman et al., 2022) but none confirmed that selection resulted in expression changes nor that AFC were linked with individual fitness variation or phenotypic changes.

In contrast, our joined quantitative and population genetics approach allowed us to show that AFC parallelism of expressed genes was positively associated with the intensity of selection on their expression levels. However, the genes showing evolutionary changes in gene expression were not associated with stronger AFC parallelism at the SNPs they carried, although expression changes were predicted by genetic selection. We could thus not link sequence and expression evolution within genes. Similarly, all eQTLs were associated with expression variation at other genes (*trans* effects), but not with local expression, and eQTL carrier genes were enriched in AFC parallelism. These results indicate that selection acts on two levels of variation: variation within expressed sequences linked to expression regulation of other genes and variation in expression levels. They also surely indicate presence of selection on regulatory variation outside expressed sequences that we do not capture with our RNA-seq approach. An additional explanation for a lack of change in expression levels despite allele frequency changes could be a change in gene product properties that is expression independent, as for instance a change of enzyme activity or affinity favoured by selection in the new environment (*e.g.*, Brown et al., 2009). Future DNA sequencing should allow us to better understand the importance of regulatory variation as well as selection on allele specific expression levels.

Nevertheless, hub genes, with the highest connectivity, showed the highest amount of indirect selection, and were enriched in evolutionary changes at the sequence (AFC parallelism) and phenotypic levels. They were also enriched in variants associated with expres-

sion variation (eQTL) at other genes. In addition, the eQTLs with higher pleiotropy, *i.e.*, influencing a larger number of expression phenotypes, were also under stronger selection. This shows how indirect selection induced evolutionary changes and confirms that selection primarily acts through indirect effects by targeting highly connected (pleiotropic) genes potentially affecting expression levels of many genes in expression modules.

Conclusions

From our results, we can conclude that more connected, and hence more pleiotropic genes are likely under stronger selection during adaptation to heat and drought stress in *T. castaneum* and thus display larger evolutionary changes at the sequence and expression levels. Similar results have been obtained in very different organisms and environments, ranging from bacteria in the laboratory (Ruelens et al., 2023) to trees adapting to environmental conditions in natural populations (Whiting et al., 2024). They point to the importance for (local) adaptation of larger effects on fitness or phenotypes that are possible at pleiotropic genes. While deleterious under stabilizing selection, those effects are beneficial during adaptation to new environments in micro-organisms (McGee et al., 2016; Maddamsetti et al., 2017; Ruelens et al., 2023) and more complex organisms (Rennison and Peichel, 2022; Thorhölludottir et al., 2023) or favored during adaptation with gene-flow in trees (Whiting et al., 2024). It thus emerges that pleiotropy and here centrality of genes in gene co-expression networks play a fundamental, positive role in the process of adaptation. Moreover, our study demonstrates that gene expression evolution is predictable during rapid adaptation and driven by indirect selection pressures on correlated expression phenotypes.

Through the integration of gene expression, fitness, and allele frequency data, we gained insights into the mechanisms underlying evolutionary changes during rapid adaptation. Our findings emphasize the importance of measuring selection acting on gene expression variation and the critical information on the molecular basis of adaptation that can be gained from such measurements. Our study highlights the importance of considering genetic correlations and network structures for predicting and understanding gene expression evolution. More similar studies would confirm the generality of our findings.

Acknowledgments

We thank Maria Domenica Moccia, Elke Karaus, Martina Berchtold, Encarnación Lozano for RNA extraction and quality control, Lucy Poveda and Maria Domenica Moccia for advice in library preparation and sequencing, Lennart Opitz for helping to process the RNA-seq data. We also thank Outi Savolainen, Tanja Pyhäjärvi, Chedly Kastally and Pierre Nouhaud for helpful suggestions and discussions, and Hannu Mäkinen (BioData Analytics unit, University of Helsinki) for his help with WGCNA analyses.

Funding

E.L.K, C.B.R. and F.G. were supported by the Swiss National Science Foundation, grants PP00P3_144846 and PP00P3_176965 to F.G.

Conflicts of interest

The authors declare that there is no conflict of interest.

Data accessibility

All the supplementary data and scripts, including the complete dataset for the genomics analysis, are available on Zenodo at https://doi.org/10.5281/zenodo.11080749. The complete transcriptomics pipeline and analysis scripts related to the genomics analysis are freely available on Github at https://github.com/charlesrocabert/Koch-et-al-Predictability-of-Gene-Expression. Sequence reads of the G1 samples are available on GEO (90 samples) and ENA (396 samples) platforms with accession numbers GSE133570 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133570) and PRJEB76655 (https://www.ebi.ac.uk/ena/browser/view/PRJEB76655) respectively. Sequence reads of the G21 samples are accessible at GEO, with accession number GSE156256 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15

6256). Sample accession numbers and description are provided in a supplementary data table on the Zenodo repository.

Contributions

E.L.K. and F.G. designed research; E.L.K., C.R., C.B.R., and F.G. performed research; E.L.K., C.R., and C.B.R. analyzed data; and E.L.K., C.R., and F.G. wrote the paper.

Material and Methods

We used the *Tribolium castaneum* Cro1 strain (Milutinović et al., 2013), collected from a wild population in 2010 and adapted to lab standard conditions (33°C, 70% relative humidity) for more than 20 generations. Beetles were kept in 24h darkness on organic wheat flour mixed with 10% organic baker's yeast. We sterilized the flour and yeast at 80°C for 24 hours before use. At the start of the experiment (generation G1), we measured fitness in control (=standard) conditions and in a stress treatment with increased temperature (37°C) and decreased humidity (30% r. h.), the hot-dry treatment (HD). To estimate genetic variances in both conditions, we performed a split-brood paternal half-sib breeding design with 147 half-sib families from six replicate lines (see Koch and Guillaume, 2020a; Koch et al., 2020). We created the half-sib families by mating a virgin male to three virgin females at the age of four weeks, in three subsequent rounds of mating per male. This resulted in three full-sib families per male. This nested paternal half-sib-breeding design enabled us to estimate additive genetic variances and co-variances (Falconer and Mackay, 1996). Mating and egg laying of each single female took place during four days in control conditions in 15 mL tubes with 1g of medium. A second round of egg laying was performed for four days to produce a second tube for each female. Each tube was then randomly assigned to either the control (CT) or hot-dry (HD) treatment, resulting in full-sib families (F1) split between CT and HD. Four F1 females and two F1 males per full-sib family and condition were separated at the pupal stage and transferred to 10 mL tubes with 1 g of medium and remained there until they were used for the fitness assay eight weeks later.

Fitness assay in generation 1

We measured fitness as the reproductive output per female by counting the number of surviving adult offspring of 2,937 virgin females (1,518 in Control and 1,419 in Hot-Dry). We mated each virgin female with one unrelated male from the same condition in 15 mL tube with 1 g medium. The male was removed after 24 h. Females were removed from the tubes after one week of egg laying, and 9 g of medium was added to provide food for the developing offspring. Females were flash-frozen in liquid nitrogen and stored at -80°C until sequencing. The number of adult offspring was counted after five weeks.

Experimental evolution

We used seven replicate lines per condition originating from the Cro1 ancestral population and used for the fitness assay in G1. We let each line adapt for 20 non-overlapping generations in CT and HD. Each line consisted of 120 breeding individuals. Each new generation was set by randomly selecting 120 pupae and placing them into a new vial with 80g of fresh medium.

Reciprocal transplant after 20 generations

At generation 21, we performed a common garden reciprocal transplant experiment to assess adaptation by measuring fitness as in G1 (see Koch and Guillaume, 2020b). Adult beetles from all selection lines in CT and HD were transferred to CT conditions, in which they stayed for one generation before creating full-sib families within each line from both CT and HD origins. Each full-sib family was then split again between CT and HD at the egg stage (generation 22). With this design, we could assess fitness differences of full-sib families in both conditions from lines originated from either CT or HD. Fitness was assessed by mating a virgin male and a virgin female of the same selection line, but from different families. After four days of mating and egg laying, males and females were removed from the medium and flash-frozen in liquid nitrogen to be stored at -80°C until sequencing. The number of adult offspring produced within four days was counted after four weeks in CT and five weeks in HD to account for a difference in development time.

Sampling, RNA extraction, library preparation, sequencing and data processing

RNA-seq was performed on 486 female beetles with known fitness at generation 1 (244 from CT and 242 from HD) and on 158 females with known fitness at generation 21 (90 from CT and 68 from HD). Generation 1 samples originated from a half-sib/full-sib cross (see above) with approximately 21 sires, 3 dams/sires and 8 full-sib females/dam, which were separately grown in CT and HD in batch of 4 females per full-sib family. In generation 21, we sampled four females per family per environment from seven full-sib families and lines.

The same extraction and library preparation protocols were followed for each set of individuals sequenced separately at generation 1 and 21. Single individuals were homogenized in Tri-Reagent® (Zymo Research, California, USA) using an electric bead mill. RNA was extracted with the RNA Mini Prep kit (Zimo Research, California, USA) following the instructions of the manufacturer. RNA-quality was checked with the Bioanalyzer 2100 (Agilent, Waldbronn, Germany) and concentrations were measured with a Qubit® Fluorometer (Life Technologies, California, USA). Libraries were created with 500 ng RNA for each individual separately with the LEXOGEN mRNA-Seq Library Kit following the manual (LEXOGEN GmbH, Vienne, Austria). Library quality was checked on a TapeStation (Agilent, Waldbronn, Germany) and concentrations were determined by qPCR. Libraries were diluted to the same molarity. Concentration of dilutions were checked again by qPCR and libraries were pooled (36 libraries per pool). The treatments were randomized during RNA-extraction, library preparation, and sequencing. The single-end sequencing was performed in five runs on the Illumina NextSeq 500 (Illumina, Inc, California, USA) using the 75 cycles High Output Kit. Each run resulted in 550-600 million reads that passed the internal sequencer filter. After quality control using FastQC (`www.bioinformatics.bbsrc.ac.uk/projects/fastqc`), reads were mapped against the reference genome (`ftp://ftp.ensemblgenomes.org/pub/release30/metazoa/gtf/tribolium_castaneum/Tribolium_castaneum.Tcas3.30.gtf.gz`) with STAR v.2.5 (Dobin et al., 2013) (adapters were trimmed and the first 10 bases were hard trimmed, minimum average quality Q10, minimum tail quality 10, minimum read length 20). We then used FeatureCounts (Liao et al., 2014) to count the number of reads that mapped

to each gene in the reference genome. Mapping as well as read counting was performed within the data analysis framework SUSHI (Hatakeyama et al., 2016). We normalized read counts to count-per-million (cpm) with the R package edgeR (Robinson et al., 2010) using TMM normalization. Only genes with a minimum of 1 cpm in at least two samples were kept. We did this for each condition separately resulting in slightly different gene numbers in CT and HD at generation 1 (CT: 11,812 genes, HD: 11,890 genes) and 11,527 genes for generation 21.

RNA-seq single nucleotide polymorphism detection

We used GATK-4.2.3.0 (Van der Auwera and O'Connor, 2020) to detect genomic variants from a total set of 822 whole transcriptome RNA-seq samples, following GATK best practices and previous approaches (Adetunji et al., 2019; Jehl et al., 2021). Individuals raised in different conditions at G21 (see Koch and Guillaume, 2020b) were added to the set of 486 samples to improve genotype imputation (see below). Mapped RNA-sequences (BAM files) were edited to add read groups (samtools 1.14, Danecek et al. 2021), and MAPQ values equal to 255 were re-calibrated to 60 to match GATK format. Duplicated reads were marked (picard Duplicates function, Broad 2019), and N-cigar reads were split using GATK SplitNCigarReads function. Variant detection was done for each sample with GATK HaplotypeCaller using the parameters -ERC GVCF –pcr-indel-model NONE –dont-use-soft-clipped-bases. The joint variant call was performed using GATK GenomicsDBImport and GenotypesGVCFs functions with parameter –max-genotype-count 1500. Bi-allelic single nucleotide polymorphisms (SNPs) were then selected using GATK SelectVariants with parameters –select-type-to-include SNP –restrict-alleles-to BIALLELIC –exclude-non-variants. Missing genotypes (DP=0) were labeled as missing ("./.") using bcftools-1.14 (Danecek et al., 2021), and SNP tags were re-calculated using bcftools +fill-tags function. We then removed low-quality SNPs using GATK VariantFiltration function with six filters: QD $<$ 2, FS $>$ 30, SOR $<$ 40, MQ $<$ 40, MQRankSum $<$ -12.5 and ReadPosRankSum $<$ -8.0. SNPs were annotated on the reference genome using snpEff-5.1 software (Cingolani et al. 2012, default interval size of 5kb) resulting in 566,296 quality-checked SNPs (Data S2). Finally, we filtered out SNPs with a call rate lower than 50% using GATK VariantFiltration function, leading to a VCF file containing 358,142

SNPs (online supporting file Data S3).

Imputation of missing genotypes

Around 38% of the genotypes are missing in the dataset, compromising non-missing markers' overlap between generations in each line. To impute missing genotypes, we used Beagle-5.4 software (Browning et al., 2021) with the VCF file as a sole input. Imputed markers were then merged with the original VCF file and SNP tags were re-calculated using bcftools *+fill-tags* function (files available online in Data S4). Importantly, we tested Beagle-5.4 accuracy by building a toy dataset from all SNPs with a call rate of 100% (7,844 SNPs), and by generating missing markers using the same distribution as in the complete VCF file. In all cases, more than 98% of genotypes were accurately imputed (see Appendix S1). Imputed data available online as Data S5).

eQTL associations

We used the imputed VCF file to calculate eQTL associations from the half-sib/full-sib cross at generation 1 in HD. SNPs with a minor allele frequency (MAF) lower than 0.05 were filtered out, yielding a set of 141,387 SNPs (see Data S6). After applying a quantile normalization on the expression of each gene (reference: standard normal distribution), genome-wide association analyses (GWAA) were performed on a gene-wise basis to detect eQTLs for each gene expression level as a phenotype by using a linear mixed model implemented in the software GEMMA-0.98.5 (option -lmm; Zhou and Stephens 2012). GEMMA was run with the SNP-based kinship matrix (genomic relatedness matrix) as random factor. The kinship matrix was calculated with GEMMA prior to running the association analysis. As GEMMA does not accept VCF files in input, the VCF file was converted into a .bed file with PLINK2 (Chang et al., 2015). Phenotypes (gene expression data) were further added to a .fam file (PLINK2) containing the pedigree structure of the data. For each GWAA, a GC correction was applied on p-values (Bacanu et al., 2000), and a FDR of 5% was applied to detect and collect significant eQTL associations. We found 958 eQTLs significantly associated to 556 gene expression levels. The same analysis was run by using relative fitness measurements as a phenotype, no significant associations were found. Results available online as Data S7.

656 ***Allelic frequency changes***

657 We used the imputed VCF file to perform the analysis of allele frequency changes (AFCs)
658 within lines between generations 1 and 21, in CT and HD separately. In CT and HD,
659 four lines (L1, L3, L5, L6) were sequenced at generations 1 and 21. Additional lines were
660 sequenced at G21 only: (L4, Mx1, and Mx3) in CT and (L2, Mx1, and Mx2) in HD. To
661 calculate AFCs, we used allele frequencies at generation 1 when available (L1, L3, L5, L6),
662 or we used a common pool at generation 1 (all available lines merged in a single pseudo-
663 population). To do so, we split the imputed VCF file into one file per line and generation
664 for each environment, as well as two additional files with all samples at generation 1 for
665 CT and HD. SNP tags were then re-calculated using bcftools +fill-tags function. Allele
666 frequencies were polarized by MAF at generation 1 to match the eQTL analysis. For each
667 line and each SNP, the allelic frequency change (AFC) was then the difference between
668 allele frequencies at generation 21 and 1. The resulting dataset is available online as Data
669 S8.

670 ***Detection of significant allele frequency changes and parallelism***

671 To detect significant allele frequency shifts in each line of HD environment, we used the
672 distribution of AFCs in CT environment as reference distribution for a non-parametric
673 bootstrap procedure. We tested whether the tail probability of each SNP's AFC within
674 each HD line at G21 was larger than 97.5% based on its confidence interval derived from
675 a sample of 10,000 bootstrapped AFC distributions from the full AFC distribution in CT
676 (*i.e.*, from the pooled lines). The number of AFCs randomly sampled in each bootstrap
677 replicate equated the number of SNPs within each HD line. This allowed us to test for
678 significant AFC in all seven HD lines sequenced in G21. We then defined AFC parallelism
679 as the number of lines in which a SNP was showing a significant AFC. Parallelism thus
680 varies between 0 and 7. The full results are available online as Data S9.

681 ***Differential gene expression analysis***

682 Evolutionary responses of gene expression after 20 generations of adaptation in HD were
683 measured and analyzed in Koch and Guillaume (2020b). Briefly, we compared gene
684 expression between CT and HD lines in HD condition in the reciprocal transplant at

G21 described above. Expression was measured by RNA-seq in 56 individuals from seven replicate CT lines and seven replicate HD lines. Differential gene expression was assessed with normalized and standardized read counts with the R package *limma* (Law et al., 2014; Ritchie et al., 2015). We used the *duplicatecorrelation* function in *limma* to account for non-independence between samples from the same line (*i.e.*, sequencing batch and maternal effects). See Koch and Guillaume (2020b) for more details. This analysis yielded 55 genes with a significant evolutionary response of their transcript abundance in HD at generation 21 after correction for multiple testing (Koch and Guillaume, 2020b). From the same study, we retrieved the complete set of up- and down-regulated genes in HD lines compared to their counterparts in CT lines in HD.

Quantitative Genetics analyses of selection on gene expression

We estimated genetic selection acting on gene expression from the genetic covariance of the trait with relative fitness $s_g = cov_A(z, w)$ (Robertson, 1966; Price, 1970) in 486 individuals sampled at the first generation of the experiment. We implemented a Bayesian sparse factor analysis of genetic covariance matrices (BSFG) method as described in Runcie and Mukherjee (2013). BSFG uses a linear mixed model and sparse factorization to estimate the additive genetic covariance matrix $\mathbf{G}$ and the residual covariance matrix $\mathbf{R}$ among traits. We utilized BSFG to estimate the genetic covariance between the expression level of each gene and relative fitness and between pairs of genes. To run BSFG, we used the MATLAB (MATLAB version 2014b) code provided at `www2.stat.duke.edu/~sayan/bfgr/Instructions/soft.html`. We included a batch effect (sequencing runs) as fixed factor and mother identity as random factor to account for potential maternal and common environment effects. Since it was not possible to analyze the whole data in one single analysis, we divided the whole data set randomly into ten smaller subsets (each consisting of 1181 genes in CT, 1189 genes in HD). We repeated this subdivision five times, resulting in five subdivisions each consisting of ten different subsets. Relative fitness was included as a trait in all subsets. All of these subsets were analyzed three times (three replicates) with 50,000 iterations of BSFG each. The first 20,000 iterations were discarded as burn-in and the remaining chain thinned for every 30th sample resulting in 1000 posterior samples for each replicate. Convergence was assessed by using Gelman-Rubinstein diagnostics $\hat{R}$

(Gelman and Rubin, 2007; Brooks and Gelman, 1998) for the three replicates. MATLAB output was further analyzed in R (version 3.4.2, R Core Team, 2017). The posterior distribution of $\mathbf{G}$ was calculated from the output of BSFG as described in Runcie and Mukherjee (2013). The genetic correlation of transcript expression with fitness (r_g) was computed as $r_g = \frac{s_g}{\sqrt{V_{A_w} \cdot V_{A_z}}}$, where V_{A_w} and V_{A_z} are the additive genetic variances of relative fitness and of the transcript expression level, respectively. Additive genetic variances correspond to the diagonal elements of $\mathbf{G}$. The off-diagonal elements hold the pairwise genetic covariance between gene expression levels. We further decomposed genetic selection s_g into its direct and indirect selection components following Stinchcombe et al. (2014). Accordingly, we computed the gradient of direct genetic selection on gene expression levels as $\beta_g = \mathbf{G}^{-1} s_g$ from the posterior distributions of $\mathbf{G}$ and s_g. As s_g is the sum of direct and indirect selection gradients, the difference $(s_g - \beta_g)$ indicates how selection acting on correlated genes may affect evolutionary changes in gene expression at a given gene. Constraints can be indicated if the predicted response s_g is opposite to the direct selection acting on a gene expression level. Another sign of constraints can be a smaller total predicted response compared to the direct selection. We compared the direction of s_g and β_g and quantified the proportion of the predicted response that was due to direct selection as β_g / s_g.

We obtained five estimates of s_g and β_g for each gene, one for each subdivision of the dataset. Each estimate and its significance are evaluated from the pooled posterior distributions of the three independent runs of BSFG. The estimate's value is the mode of the joint posterior and significance is reached if the 95% HDI does not overlap with zero. We report the average mode of s_g and β_g over the five subdivisions for genes with at least one significant value in HD. We used the R package *MCMCglmm* (Hadfield, 2010) to compute the posterior modes, highest density intervals (HDI).

Weighted gene co-expression network analysis (WGCNA)

To get more insights into the genetic architecture of transcriptomes, we applied a weighted gene co-expression network analysis (WGCNA) as implemented in the R package *WGCNA* (Langfelder and Horvath, 2008). We constructed gene co-expression networks for CT and HD conditions separately at generation 1 and 21. At generation 1, we used the same

data set as for the Bayesian sparse factor analysis (see main text). We applied the same general approach to data sets at G1 and G21. Prior to analysis, lowly expressed genes (minimum of 1 cpm in at least two samples) were removed. We then normalized read counts (cpm, TMM-normalized) and corrected for families and sequencing runs using the *removeBatchEffect* function in the limma R package (Pritchard et al., 2017). Briefly, WGCNA constructs networks using the absolute value of the Pearson's correlation coefficient as the gene co-expression measure, which is raised to a power β to create the adjacency matrix. The topological overlap distance calculated from the adjacency matrix is then clustered with the average linkage hierarchical clustering. For choosing power β we followed the recommendations of (Zhang and Horvath, 2005) and used the scale-free topology criterion resulting in $\beta = 4$ in Control, $\beta = 5$ in Hot-Dry at G1. We used the *blockwiseModule* function in the R-package WGCNA to construct a signed network using a minimum module size of 20 and a *mergeCutHeight* of 0.1 (threshold for merging modules with highly correlated module eigengenes). Other parameters remained at their default. To test whether a module was related to offspring number, the correlation between a module eigengene (corresponds to the first principal component) and offspring number was used. Module preservation statistics were computed using the *modulePreservation* function (1000 permutations) implemented in WGCNA (Langfelder et al., 2011). Network module preservation statistics quantify how density and connectivity patterns of modules defined in a reference data set are preserved in a test data set without the need to define modules in the test data set. The composite measure *Zsummary* and *medianRank* were used to evaluate module preservation in other conditions. *Zsummary* is a composite statistic that combines preservation of density as well as connectivity of a module. Since *Zsummary* strongly depends on module size, it is only appropriate to examine preservation of the same module in different conditions. Values below two indicate no preservation, values between two and ten moderate and above ten strong preservation. To compare different modules within the same condition *medianRank* (Figure S2) is more informative which gives relative preservation of modules (Langfelder et al., 2011). For hub gene identification we used the *intramodularConnectivity* function of the WGCNA R package and selected the top ten genes of each model showing the highest intramodular connectivity. We then tested whether these hub genes were enriched in eQTLs, and whether they

significantly differed from other genes in selection and evolutionary responses.

References

Adetunji, M. O., Lamont, S. J., Abasht, B., and Schmidt, C. J. (2019). Variant analysis pipeline for accurate detection of genomic variants from transcriptome sequencing data. PLOS ONE, 14(9):1–17.

Agrawal, A. F. and Stinchcombe, J. R. (2009). How much do genetic covariances alter the rate of adaptation? Proc. R. Soc. B, 276(1659):1183–1191.

Ahmad, F., Debes, P. V., Nousiainen, I., Kahar, S., Pukk, L., Gross, R., Ozerov, M., and Vasemägi, A. (2021). The strength and form of natural selection on transcript abundance in the wild. Molecular Ecology, 30(12):2724–2737.

Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C., and Jones, A. G. (2008). Understanding the evolution and stability of the G-matrix. Evolution, 62(10):2451–2461.

Ayroles, J. F., Carbone, M. A., Stone, E. A., Jordan, K. W., Lyman, R. F., Magwire, M. M., Rollmann, S. M., Duncan, L. H., Lawrence, F., Anholt, R. R., and Mackay, T. F. (2009). Systems genetics of complex traits in *Drosophila melanogaster*. Nature Genetics, 41:299–307.

Bacanu, S.-A., Devlin, B., and Roeder, K. (2000). The power of genomic control. The American Journal of Human Genetics, 66(6):1933–1944.

Barabási, A. L. and Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. Nature Reviews Genetics, 5:101–113.

Barghi, N., Hermisson, J., and Schlötterer, C. (2020). Polygenic adaptation: a unifying framework to understand positive selection. Nature Reviews Genetics, 21(12):769–781.

Barghi, N., Tobler, R., Nolte, V., Jaksić, A. M., Mallard, F., Otte, K. A., Dolezal, M., Taus, T., Kofler, R., and Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in Drosophila. PLOS Biology, 17(2):e3000128.

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., and Petrov, D. A. (2014). Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. PLOS Genet., 10(11):e1004775.

Blows, M. W., Allen, S. L., Collet, J. M., Chenoweth, S. F., and McGuigan, K. (2015). The phenome-wide distribution of genetic variance. Am. Nat., 186(1):15–30.

Blows, M. W. and Hoffmann, A. A. (2005). A reassessment of genetic limits to evolutionary change. Ecology, 86(6):1371–1384.

Bonnet, T., Wandeler, P., Camenisch, G., and Postma, E. (2017). Bigger is fitter? quantitative genetic decomposition of selection reveals an adaptive evolutionary decline of body mass in a wild rodent population. PLOS Biology, 15(1):e1002592.

Brem, R. B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. Science, 296(5568):752–755.

Broad (2019). Picard Toolkit, Broad Institute. http://broadinstitute.github.io/picard.

Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics, 7:434–455.

Brown, K. M., DePristo, M. A., Weinreich, D. M., and Hartl, D. L. (2009). Temporal constraints on the incorporation of regulatory mutants in evolutionary pathways. Mol. Biol. Evol., 26(11):2455–2462.

Browning, B. L., Tian, X., Zhou, Y., and Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. The American Journal of Human Genetics, 108(10):1880–1890.

Burban, E., Tenaillon, M. I., and Le Rouzic, A. (2021). Gene network simulations provide testable predictions for the molecular domestication syndrome. Genetics, 220(2):iyab214.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience, 4(1). s13742-015-0047-8.

Chebib, J. and Guillaume, F. (2021). Pleiotropy or linkage? Their relative contributions to the genetic correlation of quantitative traits and detection by multitrait GWA studies. Genetics, 219(4):iyab159.

Chebib, J. and Guillaume, F. (2022). The relative impact of evolving pleiotropy and mutational correlation on trait divergence. Genetics, 220(1):iyab205.

Chevin, L.-M., Martin, G., and Lenormand, T. (2010). Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogenous mutation, and parallel evolution. Evolution, 64(11):3213–3231.

Cingolani, P., Platts, A., Coon, M., Nguyen, T., Wang, L., Land, S., Lu, X., and Ruden, D. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. Fly, 6(2):80–92.

Collet, J. M., McGuigan, K., Allen, S. L., Chenoweth, S. F., and Blows, M. W. (2018). Mutational Pleiotropy and the Strength of Stabilizing Selection Within and Between Functional Modules of Gene Expression. Genetics, 208(4):1601–1616.

Daimon, T. and Shinoda, T. (2013). Function, diversity, and application of insect juvenile hormone epoxidases (CYP15). Biotechnology and Applied Biochemistry, 60(1):82–91.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10(2). giab008.

des Marais, D. L., Guerrero, R. F., Lasky, J. R., and Scarpino, S. V. (2017). Topological features of a gene co-expression network predict patterns of natural diversity in environmental response. Proceedings of the Royal Society B-Biological Sciences, 284(1856).

des Marais, D. L., Hernandez, K. M., and Juenger, T. E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. Annual Review of Ecology, Evolution, and Systematics, 44(Volume 44, 2013):5–29. Publisher: Annual Reviews.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). Star: Ultrafast universal rna-seq aligner. Bioinformatics, 29:15–21.

Falconer, D. S. and Mackay, T. F. (1996). Introduction to Quantitative Genetics. Longman, London, fourth edition.

Fay, J. C. and Wittkopp, P. J. (2008). Evaluating the role of natural selection in the evolution of gene regulation. Heredity, 100(2):191–199.

Fisher, R. A. (1930). The genetical theory of natural selection. Oxford: Clarendon Press.

Flatt, T., Tu, M.-P., and Tatar, M. (2005). Hormonal pleiotropy and the juvenile hormone regulation of drosophila development and life history. BioEssays, 27:999–1010.

Fraser, H. B., Hirsh, A. E., Steinmetz, L. M., Scharfe, C., and Feldman, M. W. (2002). Evolutionary rate in the protein interaction networks. Science, 296(5568):750–752.

Fraser, H. B., Moses, A. M., and Schadt, E. E. (2010). Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc. Natl. Acad. Sci. U. S. A., 107(7):2977–2982.

Gelman, A. and Rubin, D. B. (2007). Inference from iterative simulation using multiple sequences. Stat. Sci., 7:457–472.

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., and Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature, 525:372–375.

Gibson, G. (2008). The environmental contribution to gene expression profiles. Nat Rev Genet, 9(8):575–581.

Glassberg, E. C., Gao, Z., Harpak, A., Lan, X., and Pritchard, J. K. (2019). Evidence for Weak Selective Constraint on Human Gene Expression. Genetics, 211(2):757–772.

Groen, S. C., Ćalić, I., Joly-Lopez, Z., Platts, A. E., Choi, J. Y., Natividad, M., Dorph, K., Mauck, W. M., Bracken, B., Cabral, C. L. U., Kumar, A., Torres, R. O., Satija, R., Vergara, G., Henry, A., Franks, S. J., and Purugganan, M. D. (2020). The strength and pattern of natural selection on gene expression in rice. Nature, 578(7796):572–576.

Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software, 33:1–22.

Hahn, M. W. and Kern, A. D. (2005). Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. Mol. Biol. Evol., 22(4):803–806.

Hansen, T. F. (2003). Is modularity necessary for evolvability? remarks on the relationship between pleiotropy and evolvability. Biosystems, 69(2-3):83–94.

Hansen, T. F. and Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. J. Evol. Biol., 21(5):1201–1219.

Hansen, T. F. and Pelabon, C. (2021). Evolvability: A quantitative-genetics perspective. Annu Rev Ecol Evol S, 52:153–175.

Hatakeyama, M., Opitz, L., Russo, G., Qi, W., Schlapbach, R., and Rehrauer, H. (2016). Sushi: An exquisite recipe for fully documented, reproducible and reusable ngs data analysis. BMC Bioinformatics, 17:1–9.

He, X. L. and Zhang, J. Z. (2006). Toward a molecular understanding of pleiotropy. Genetics, 173(4):1885–1891.

Hodgins-Davis, A., Adomas, A. B., Warringer, J., and Townsend, J. P. (2012). Abundant gene-by-environment interactions in gene expression reaction norms to copper within Saccharomyces cerevisiae. Genome Biol Evol, 4(11):1061–1079.

Houle, D. and Cheng, C. (2021). Predicting the Evolution of Sexual Dimorphism in Gene Expression. Molecular Biology and Evolution, 38(5):1847–1859.

Huang, Y. and Agrawal, A. F. (2016). Experimental evolution of gene expression and plasticity in alternative selective regimes. PLOS Genet, 12(9):e1006336.

Ihmels, J., Friedlander, G., Bergmann, S., Sarig, O., Ziv, Y., and Barkai, N. (2002). Revealing modular organization in the yeast transcriptional network. Nature Genetics, 31:370–377.

Innocenti, P. and Chenoweth, S. F. (2013). Interspecific divergence of transcription networks along lines of genetic variance in Drosophila: dimensionality, evolvability, and constraint. Mol. Biol. Evol., 30(6):1358–1367.

Jehl, F., Degalez, F., Bernard, M., Lecerf, F., Lagoutte, L., Désert, C., Coulée, M., Bouchez, O., Leroux, S., Abasht, B., Tixier-Boichard, M., Bed'hom, B., Burlot, T., Gourichon, D., Bardou, P., Acloque, H., Foissac, S., Djebali, S., Giuffra, E., Zerjal, T., Pitel, F., Klopp, C., and Lagarrigue, S. (2021). Rna-seq data for reliable snp detection and genotype calling: Interest for coding variant characterization and cis-regulation analysis by allele-specific expression in livestock species. Frontiers in Genetics, 12.

Jordan, I. K., Marino-Ramirez, L., Wolf, Y. I., and Koonin, E. V. (2004). Conservation and coevolution in the scale-free human gene coexpression network. Molecular Biology and Evolution, 21(11):2058–2070.

Josephs, E. B., Lee, Y. W., Stinchcombe, J. R., and Wright, S. I. (2015). Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression. Proceedings of the National Academy of Sciences, 112(50):15390–15395.

Josephs, E. B., Wright, S. I., Stinchcombe, J. R., and Schoen, D. J. (2017). The Relationship between Selection, Network Connectivity, and Regulatory Variation within a Population of Capsella grandiflora. Genome Biology and Evolution, 9(4):1099–1109.

Kelly, D. E., Ramdas, S., Ma, R., Rawlings-Goss, R. A., Grant, G. R., Ranciaro, A., Hirbo, J. B., Beggs, W., Yeager, M., Chanock, S., Nyambo, T. B., Omar, S. A., Woldemeskel, D., Belay, G., Li, H., Brown, C. D., and Tishkoff, S. A. (2023). The genetic and evolutionary basis of gene expression variation in east africans. Genome Biology, 24(1).

Kita, R., Venkataram, S., Zhou, Y., and Fraser, H. B. (2017). High-resolution mapping of cis-regulatory variation in budding yeast. Proc. Natl. Acad. Sci. U. S. A., 114(50):E10736–E10744.

Koch, E. L. and Guillaume, F. (2020a). Additive and mostly adaptive plastic responses of gene expression to multiple stress in *Tribolium castaneum*. PLoS Genetics, 16(5):e1008768.

Koch, E. L. and Guillaume, F. (2020b). Restoring ancestral phenotypes is a general pattern in gene expression evolution during adaptation to new environments in Tribolium castaneum. Molecular Ecology, 29(20):3938–3953.

Koch, E. L., Sbilordo, S. H., and Guillaume, F. (2020). Genetic variance in fitness and its cross-sex covariance predict adaptation during experimental evolution. Evolution, 74(12):2725–2740.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., Brotherstone, S., Guinness, F., and Clutton-Brock, T. (2002). Antler size in red deer: Heritability and selection but no evolution. Evolution, 56(8):1683–1695.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution, 33(1):402–416.

Lande, R. and Arnold, S. J. (1983). The measurement of selection on correlated characters. Evolution, 37(6):1210–1226.

Langfelder, P. and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics, 9:559.

Langfelder, P., Luo, R., Oldham, M. C., and Horvath, S. (2011). Is My Network Module Preserved and Reproducible? PLOS Computational Biology, 7(1):e1001057.

Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol., 15:R29.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7):923–930.

Lòpez-Maury, L., Marguerat, S., and Bähler, J. (2008). Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat. Rev. Genet., 9(8):583–593.

Luscombe, N. M., Madan Babu, M., Yu, H., Snyder, M., Teichmann, S. A., and Gerstein, M. (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. Nature, 431(7006):308–312.

Mack, K. L., Ballinger, M. A., Phifer-Rixey, M., and Nachman, M. W. (2018). Gene regulation underlies environmental adaptation in house mice. Genome Research, 28(11):1636–1645.

Maddamsetti, R., Hatcher, P. J., Green, A. G., Williams, B. L., Marks, D. S., and Lenski, R. E. (2017). Core Genes Evolve Rapidly in the Long-Term Evolution Experiment with Escherichia coli. Genome Biology and Evolution, 9(4):1072–1083.

Martin, G. and Lenormand, T. (2006). A general multivariate extension of fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution, 60(5):893–907.

McGee, L. W., Sackman, A. M., Morrison, A. J., Pierce, J., Anisman, J., and Rokyta, D. R. (2016). Synergistic Pleiotropy Overrides the Costs of Complexity in Viral Adaptation. Genetics, 202(1):285–295.

McGraw, E. A., Ye, Y. X. H., Foley, B., Chenoweth, S. F., Higgie, M., Hine, E., and Blows, M. W. (2011). High-dimensional variance partitioning reveals the modular genetic basis of adaptive divergence in gene expression during reproductive character displacement. Evolution, 65(11):3126–3137.

McGuigan, K., Collet, J. M., Allen, S. L., Chenoweth, S. F., and Blows, M. W. (2014). Pleiotropic mutations are subject to strong stabilizing selection. Genetics, 197(3):1051–1062.

Milutinović, B., Stolpe, C., Peuss, R., Armitage, S. A. O., and Kurtz, J. (2013). The Red Flour Beetle as a Model for Bacterial Oral Infections. PLOS ONE, 8(5):e64638.

Morrissey, M. B., Kruuk, L. E. B., and Wilson, A. J. (2010). The danger of applying the breeder's equation in observational studies of natural populations. Journal of Evolutionary Biology, 23(11):2277–2288.

Morrissey, M. B., Parker, D. J., Korsten, P., Pemberton, J. M., Kruuk, L. E., and Wilson, A. J. (2012). The prediction of adaptive evolution: empirical application of the secondary theorem of selection and comparison to the breeder's equation. Evolution, 66(8):2399–2410.

Mähler, N., Wang, J., Terebieniec, B. K., Ingvarsson, P. K., Street, N. R., and Hvidsten, T. R. (2017). Gene co-expression network connectivity is an important determinant of selective constraint. PLOS Genetics, 13(4):e1006402.

Nijhout, H. F. (1994). Insect hormones. Princeton Univ. Press, Princeton, NJ.

Orr, H. A. (2000). Adaptation and the cost of complexity. Evolution, 54(1):13–20.

Otto, S. P. (2004). Two steps forward, one step back: the pleiotropic effects of favoured alleles. Proc. R. Soc. Lond. B, 271:705–714.

Petit, A. J. R., Guez, J., and Le Rouzic, A. (2023). Correlated stabilizing selection shapes the topology of gene regulatory networks. Genetics, 224(2):iyad065.

Price, G. R. (1970). Selection and covariance. Nature, 227(5257):520–521.

Pritchard, V. L., Viitaniemi, H. M., McCairns, R. J. S., Merila, J., Nikinmaa, M., Primmer, C. R., and Leder, E. H. (2017). Regulatory architecture of gene expression variation in the threespine stickleback gasterosteus aculeatus. G3-Genes Genomes Genetics, 7(1):165–178.

Rennison, D. J. and Peichel, C. L. (2022). Pleiotropy facilitates parallel adaptation in sticklebacks. Molecular Ecology, 31(5):1476–1486.

Rifkin, S. A., Houle, D., Kim, J., and White, K. P. (2005). A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. Nature, 438(7065):220–223.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 43:e47.

Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. Animal Science, 8(1):95–108.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edger: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26:139–140.

Rudman, S. M., Greenblum, S. I., Rajpurohit, S., Betancourt, N. J., Hanna, J., Tilk, S., Yokoyama, T., Petrov, D. A., and Schmidt, P. (2022). Direct observation of adaptive tracking on ecological time scales in Drosophila. Science, 375(6586):eabj7484.

Ruelens, P., Wynands, T., and de Visser, J. A. G. M. (2023). Interaction between mutation type and gene pleiotropy drives parallel evolution in the laboratory. Philosophical Transactions of the Royal Society B: Biological Sciences, 378(1877):20220051.

Runcie, D. E. and Mukherjee, S. (2013). Dissecting high-dimensional phenotypes with bayesian sparse factor analysis of genetic covariance matrices. Genetics, 194(3):753–767.

Steige, K. A., Laenen, B., Reimegård, J., Scofield, D. G., and Slotte, T. (2017). Genomic analysis reveals major determinants of cis-regulatory variation in Capsella grandiflora. Proceedings of the National Academy of Sciences, 114(5):1087–1092.

Stern, D. L. and Orgogozo, V. (2008). The loci of evolution: how predictable is genetic evolution? Evolution, 62(9):2155–2177.

Stinchcombe, J. R., Simonsen, A. K., and Blows, M. W. (2014). Estimating Uncertainty in Multivariate Responses to Selection. Evolution, 68(4):1188–1196.

Thorhölludottir, D. A. V., Nolte, V., and Schlötterer, C. (2023). Temperature-driven gene expression evolution in natural and laboratory populations highlights the crucial role of correlated fitness effects for polygenic adaptation. Evolution, 77(9):2081–2089.

Van der Auwera, G. A. and O'Connor, B. D. (2020). Genomics in the cloud: using Docker, GATK, and WDL in Terra. O'Reilly Media.

Walsh, B. and Blows, M. W. (2009). Abundant genetic variation + strong selection = multivariate genetic constraints: A geometric view of adaptation. Annu. Rev. Ecol. Evol. Sys., 40(1):41–59.

Watson, R. A., Wagner, G. P., Pavlicev, M., Weinreich, D. M., and Mills, R. (2014). The evolution of phenotypic correlations and "developmental memory". Evolution, 68(4):1124–1138.

Whitehead, A. and Crawford, D. L. (2006). Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. U. S. A., 103(14):5425–5430.

Whiting, J. R., Booker, T. R., Rougeux, C., Lind, B. M., Singh, P., Lu, M., Huang, K., Whitlock, M. C., Aitken, S. N., Andrew, R. L., Borevitz, J. O., Bruhl, J. J., Collins, T. L., Fischer, M. C., Hodgins, K. A., Holliday, J. A., Ingvarsson, P. K., Janes, J. K., Khandaker, M., Koenig, D., Kreiner, J. M., Kremer, A., Lascoux, M., Leroy, T., Milesi, P., Murray, K. D., Pyhäjärvi, T., Rellstab, C., Rieseberg, L. H., Roux, F., Stinchcombe, J. R., Telford, I. R. H., Todesco, M., Tyrmi, J. S., Wang, B., Weigel, D., Willi, Y., Wright, S. I., Zhou, L., and Yeaman, S. (2024). The genetic architecture of repeated local adaptation to climate in distantly related plants. Nature Ecology & Evolution, pages 1–15.

Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. Nature Reviews Genetics, 8(3):206–216.

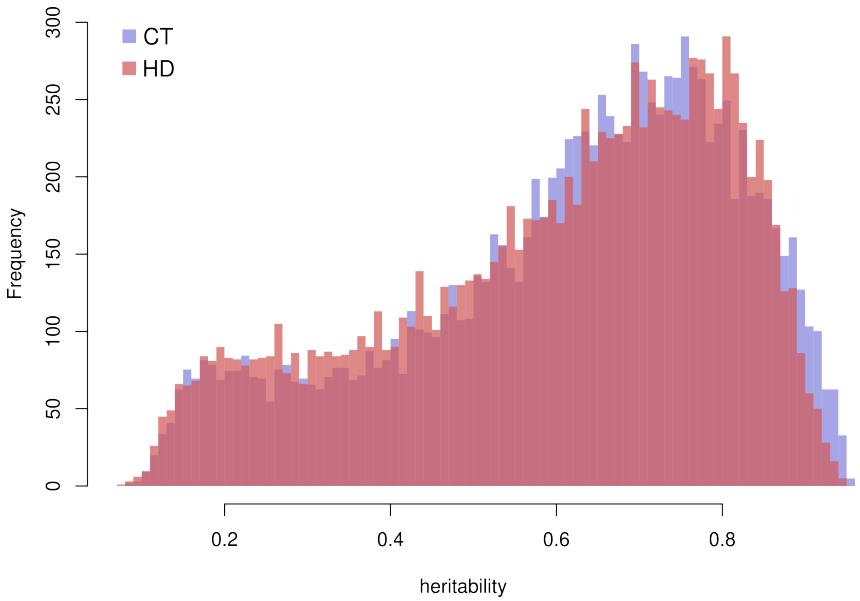
Yampolsky, L., Glazko, G. V., and Fry, J. (2012). Evolution of gene expression and expression plasticity in long-term experimental populations of drosophila melanogaster maintained under constant and variable ethanol stress. Mol. Ecol., 21:4287–4299.

Yeaman, S., Gerstein, A. C., Hodgins, K. A., and Whitlock, M. C. (2018). Quantifying how constraints limit the diversity of viable routes to adaptation. PLOS Genetics, 14(10):e1007717.

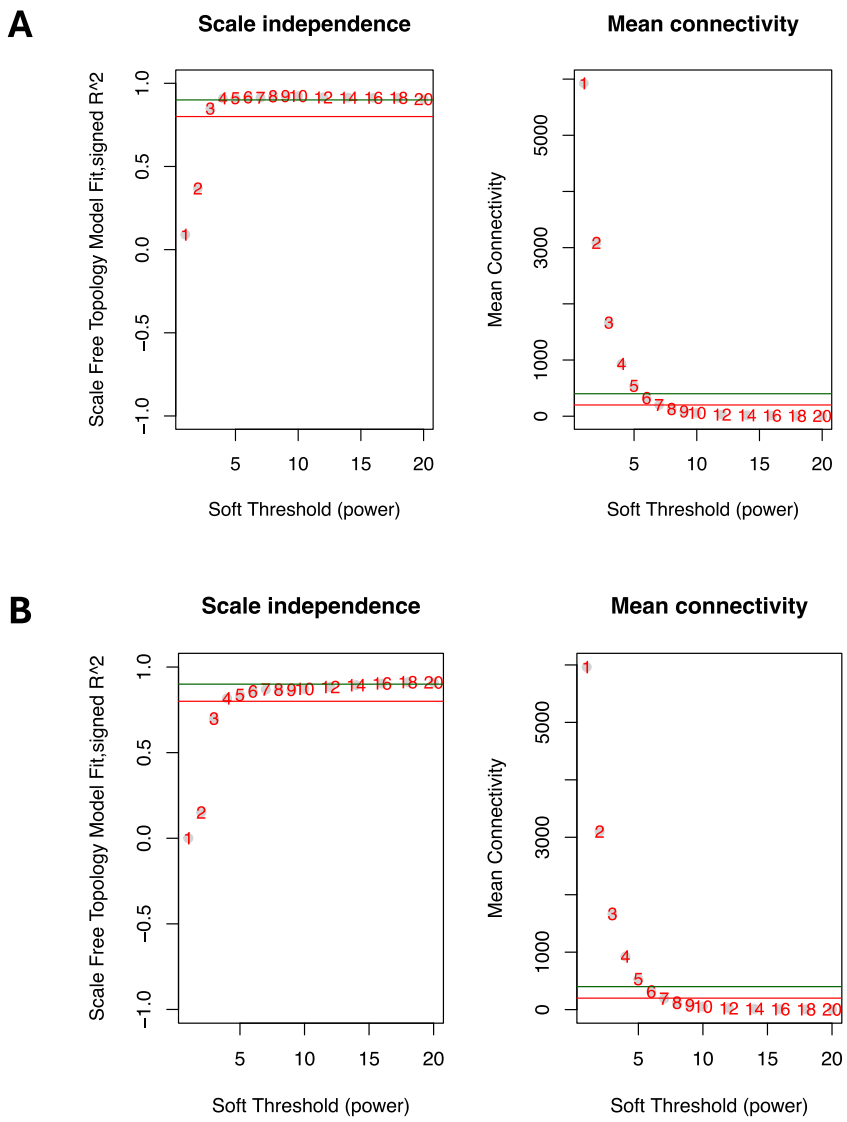
Zhang, B. and Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology, 4(1).

Zhou, X. and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nature genetics, 44(7):821–824.

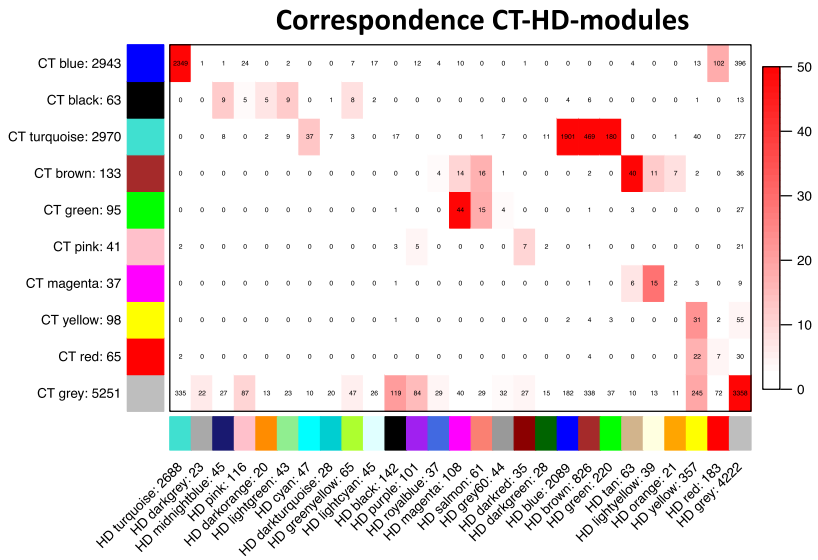
Supplementary figures



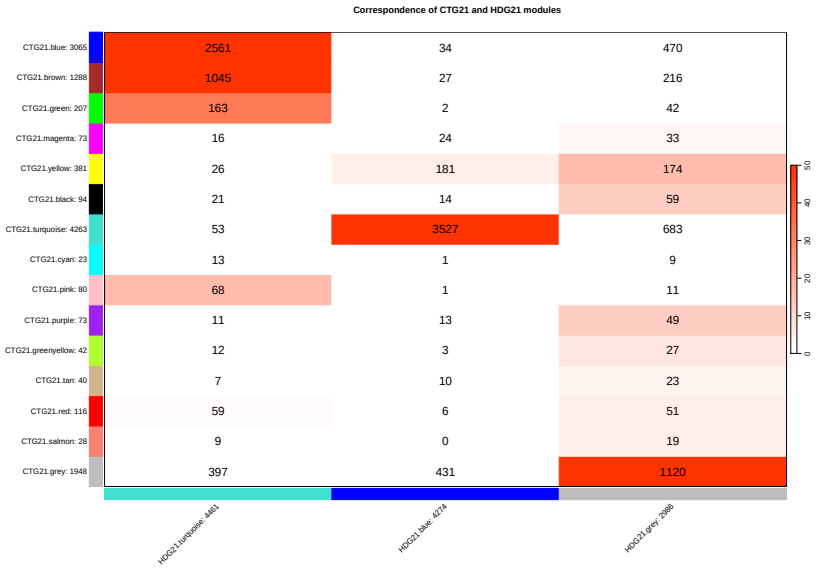
Supplementary Figure 1: Transcriptome-wide distributions of heritability h^2 of expression levels in CT (blue) and HD (red).



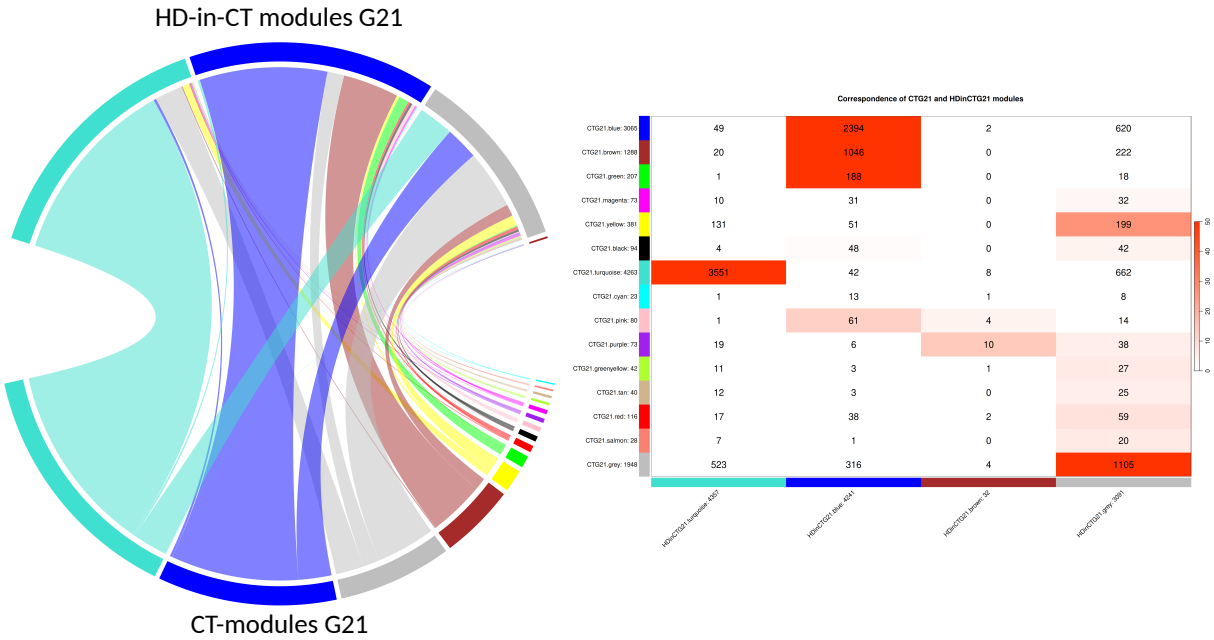
Supplementary Figure 2: Summary network indices (y-axes) as functions of the soft thresholding power (x-axes) for Control (**A** - upper row), and Hot-Dry (**B** - lower row). Numbers in the plots indicate the corresponding soft thresholding powers . The plots indicate that approximate scale-free topology is attained around the soft-thresholding power of four for Control and for six in Hot-Dry. Red line marks $R^2 = 0.8$, the green line $R^2 = 0.9$ in the scale-free topology plots. Langfelder and Horvath (2008) suggested a minimum of 0.8. Because the summary connectivity measures decline steeply with increasing soft-thresholding power, it is advantageous to choose the lowest power that satisfies the approximate scale-free topology criterion. In the 'Mean connectivity' plots, the red line indicates 200, the green line 400.



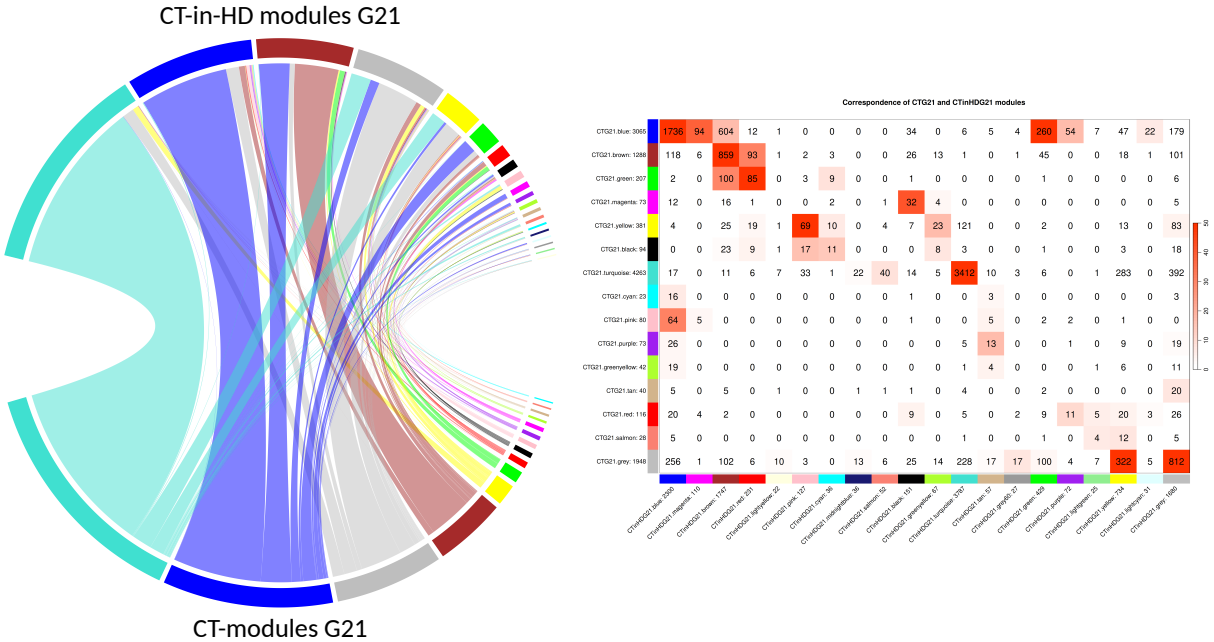
Supplementary Figure 3: Module overlap between gene co-expression networks of CT lines and HD lines at generation 1. The number of genes in each module and their overlap is shown. The intensity of red indicates significance of overlap tested with Fisher's exact test (-log10(p-value)).



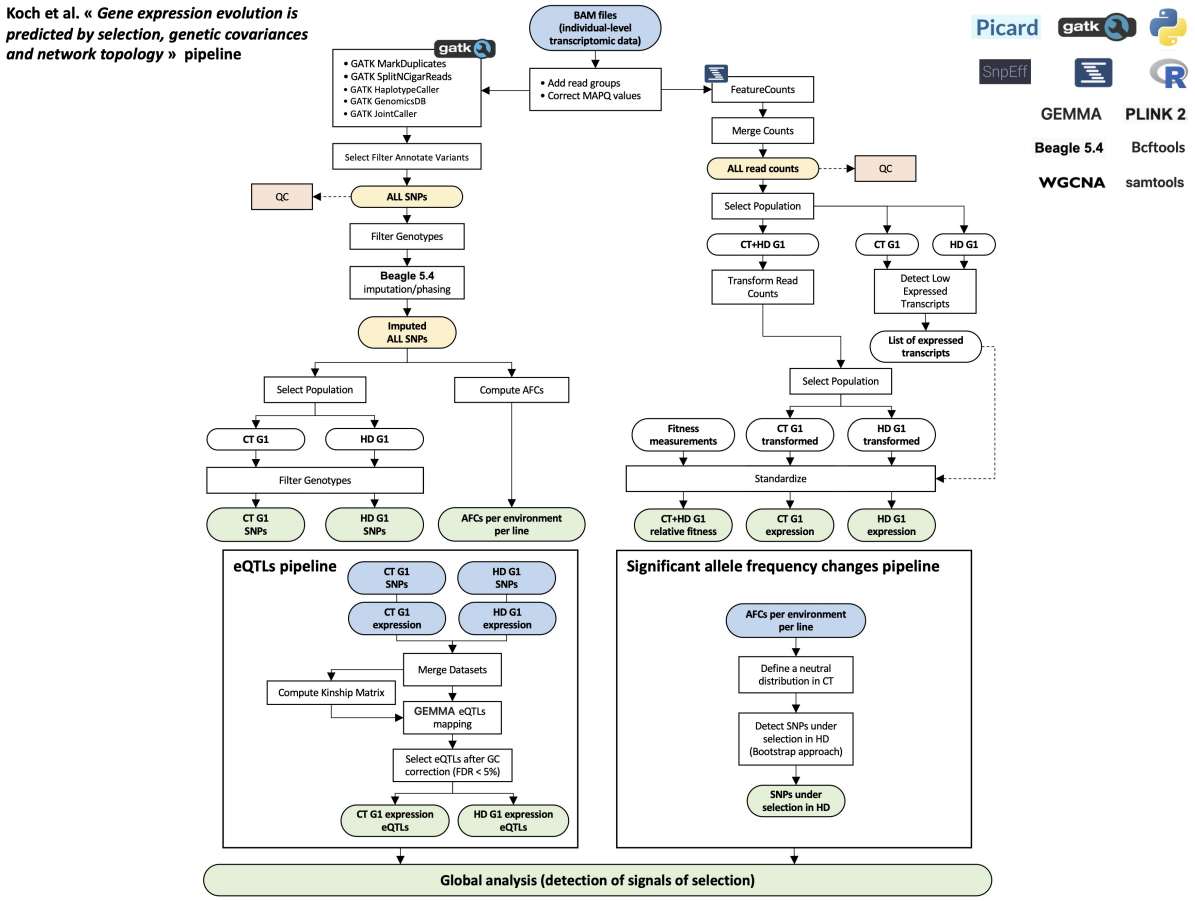
Supplementary Figure 4: Module overlap between gene co-expression networks of CT lines and HD lines at generation 21. The number of genes in each module and their overlap is shown. The intensity of red indicates significance of overlap tested with Fisher's exact test (-log10(p-value)).



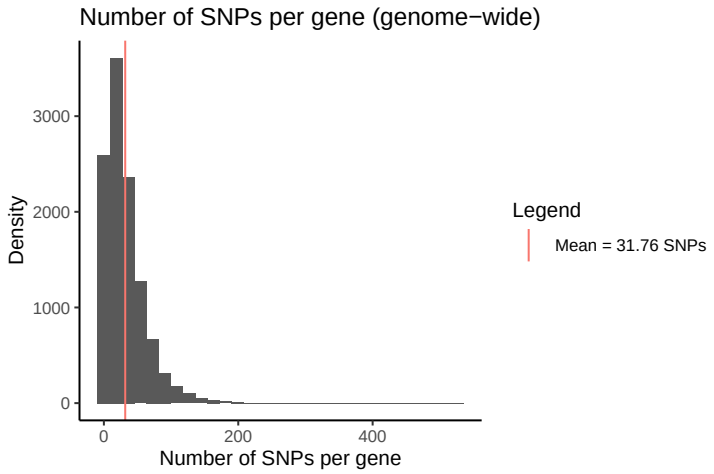
Supplementary Figure 5: Module overlap between gene co-expression networks of CT lines and HD lines who were transplanted in CT at generation 21. The number of genes in each module and their overlap is shown in the table on the right. The intensity of red indicates significance of overlap tested with Fisher's exact test (-log10(p-value)).



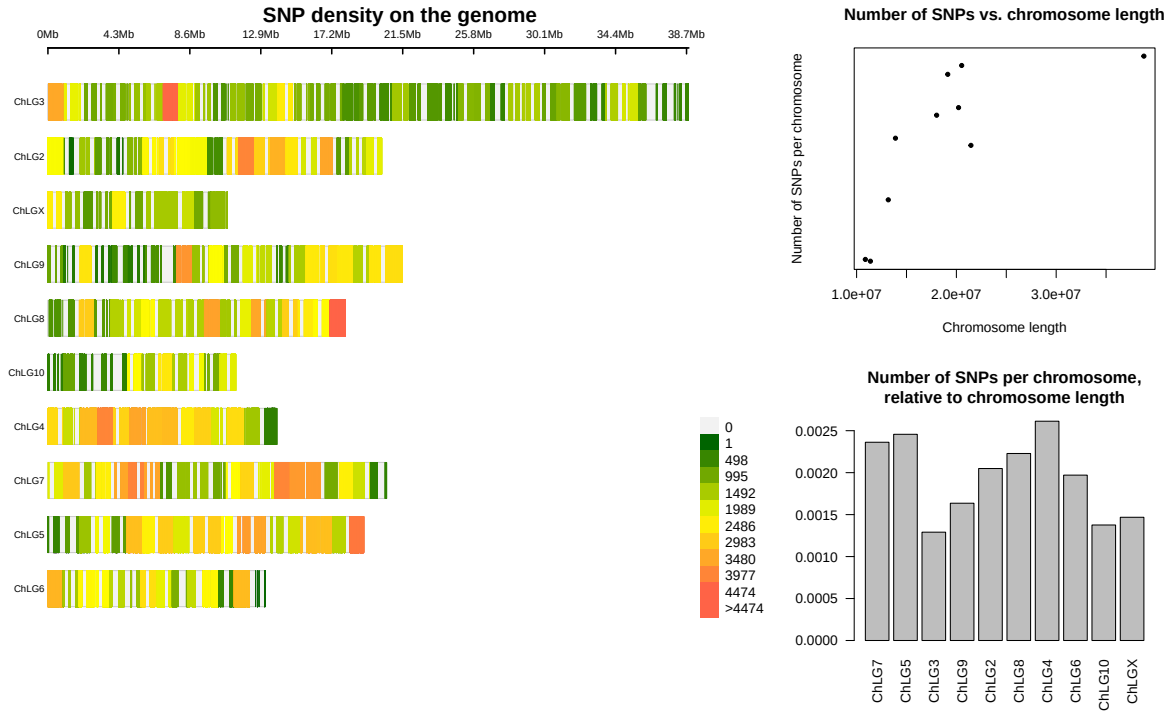
Supplementary Figure 6: Module overlap between gene co-expression networks in CT lines and CT lines who were transplanted in HD at generation 21. The number of genes in each module and their overlap is shown in the table on the right. The intensity of red indicates significance of overlap tested with Fisher's exact test (-log10(p-value)).



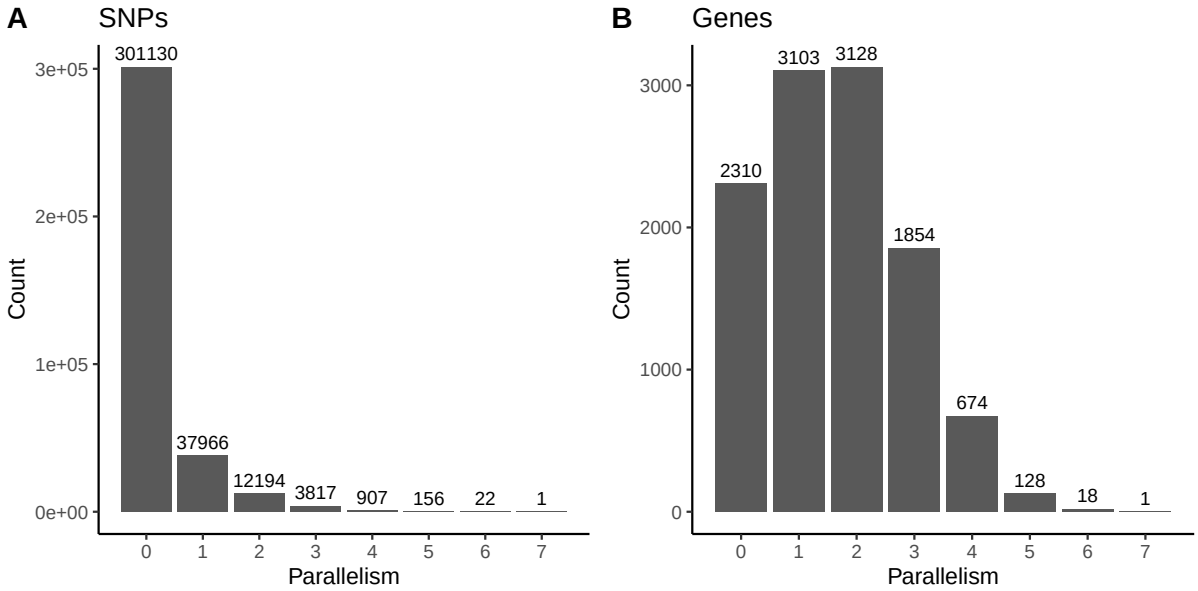
Supplementary Figure 7: Overview of the bioinformatics pipelines. The bioinformatics pipelines used to recover single nucleotide polymorphism, eQTLs and allele frequency changes (AFCs) are summarized in this diagram.



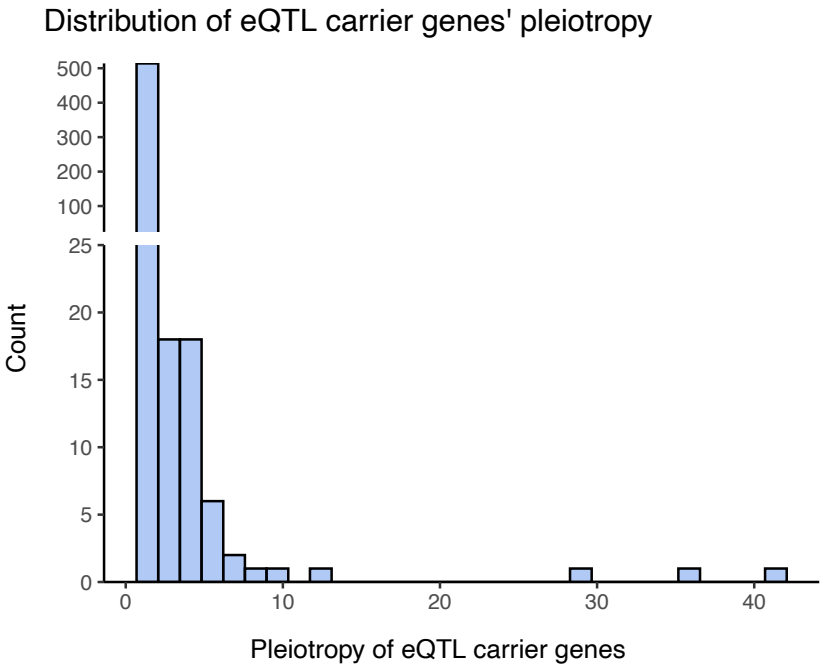
Supplementary Figure 8: Number of SNPs per gene in the genome. Histogram of the number of SNPs per gene on the entire genome. The mean number of SNPs per gene is indicated by a red vertical line (31.76 SNPs).



Supplementary Figure 9: Repartition of SNPs on the genome. Left panel: SNP density along the genome. **Top-right panel:** Number of SNPs per chromosome vs. chromosome length for the 10 chromosomes of *Tribolium castaneum*. **Bottom-right panel**: Number of SNPs per chromosome relative to chromosome length.



Supplementary Figure 10: Parallelism of SNPs and genes frequency shifts. Distribution of the parallelism of frequency shifts, from no shift (0) to fully parallel shifts in the seven lines for (**A**) SNPs, (**B**) genes. The count is shown in each category.



Supplementary Figure 11: Distributions of the pleiotropy of eQTL carrier genes. Distribution of the pleiotropy (*i.e.*, the number of significantly associated phenotypes) for all 564 eQTL carrier genes. Pleiotropy values range from 1 to 41. Vertical axis was re-scaled above $y = 25$ for readability.

Appendix S1

Gene expression evolution is predicted by selection, genetic covariances and network topology.

Appendix S1. Imputation quality testing with Beagle-5.4 software on the raw VCF file.

To test capability of Beagle-5.4 (Browning et al. 2021) to predict missing genotypes with our dataset, we used the raw VCF file obtained from the *Tribolium castaneum* transcriptomics data (see Main Manuscript) to evaluate the success rate of genotypes imputation.

To this aim, we selected in the raw VCF file all the markers having a call rate of 100% to create a subset of SNPs with no missing genotypes (7,837 markers out of 358,142, *i.e.*, ~2.19% of the initial VCF file). We then used this subset to generate 100 artificial datasets where missing genotypes were introduced at random following the same probability distribution per marker than in the initial VCF file (see Fig. 1). For each artificial dataset, we called Beagle-5.4 with the reconstructed VCF file as a sole input. For each artificial dataset, the success rate was calculated as the ratio of imputed genotypes identical to the original ones.

We found that, in mean, the success rate was ~0.98% (standard error $s.e. = 4.62 \times 10^{-6}$; see Fig. 2). In mean, Beagle-5.14 software did not impute correctly 10,915 genotypes out of 652,600 missing genotypes.

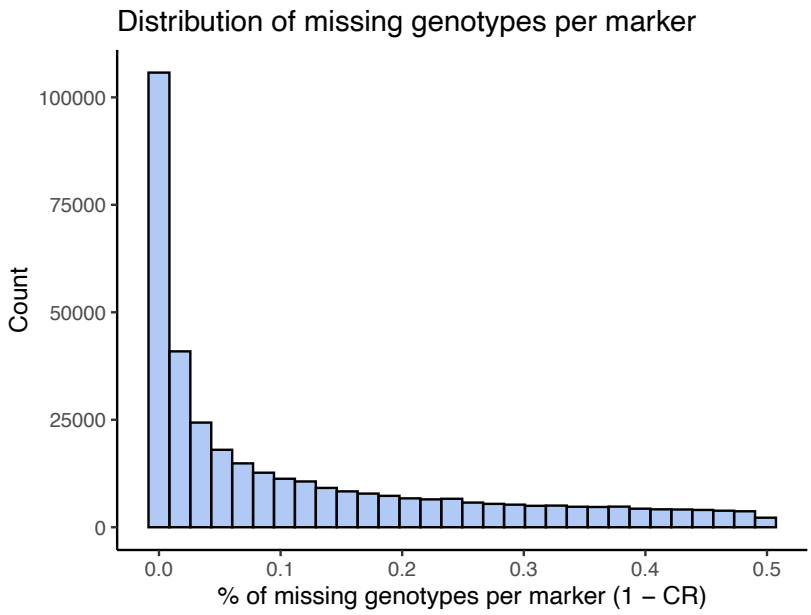


Figure 1. Distribution of the ratio of missing genotypes per marker, for the 358,142 markers present in the original VCF file. As SNPs have been filtered to keep markers with a call rate (CR) > 50% prior to this analysis, the maximum ratio of missing genotypes is 0.5.

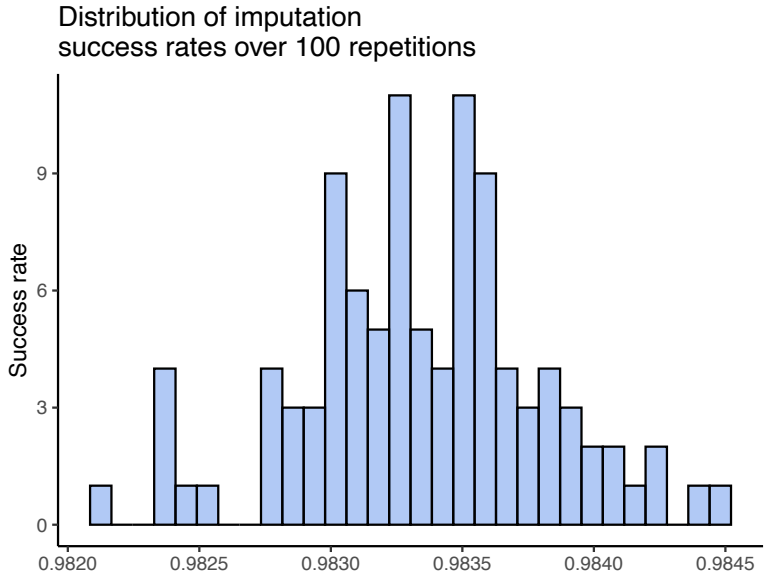


Figure 2. Distribution of the genotype imputation success rate for 100 independent tests. The success rate is the ratio of missing genotypes correctly imputed (in mean, 652,600 missing genotypes are generated by artificial dataset). Mean success rate: ~0.98% (standard error $s.e. = 4.62 \times 10^{-6}$).

References

Browning, B. L., Tian, X., Zhou, Y., and Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics*, 108(10):1880–1890.

Supporting data online

The following scripts and data sets are accessible online at https://doi.org/10.5281/zenodo.11080749.

Script S1. BSFG estimates (50.9 MB).

Script S2. WGCNA analysis and results G1 (99.6 MB).

Data S1. Datasets resulting from the global genomics analysis of the output of the transcriptomics pipeline (5.9 MB).

Data S2. VCF file containing the 566,296 quality-checked SNPs (1.1 GB).

Data S3. VCF file containing the 358,142 SNPs with a call rate higher or equal to 50% (964.3 MB).

Data S4. VCF file containing the imputed genotypes (964.3 MB).

Data S5. Results of the imputation tests (479.5 kB).

Data S6. VCF file containing imputed genotypes where SNPs with a minor allele frequency lower than 0.05 have been filtered out (MAF $\leq$ 0.05) (176.0 MB).

Data S7. List of the 1,273 significant eQTL associations (for gene expression levels or relative fitness as phenotypes) (61.2 kB).

Data S8. List of allele frequency changes (AFCs) for every markers in HD environment, for lines L1, L2, L3, L5, L6, Mx1 and Mx2 (32.4 MB).

Data S9. List of all SNPs indicating if their AFC is significantly higher in each line and their degree of parallelism (2.4 MB).