Dissecting genetic correlation through recombinant perturbations: the role of developmental bias

Haoran Cai[1], Kerry Geiler-Samerotte[2,3], David L. Des Marais[1]

[1]Department of Civil and Environmental Engineering, MIT.
15 Vassar St., Cambridge, MA, 02139 USA

[2]Center for Mechanisms of Evolution, Biodesign Institute, Arizona State University, USA

[3]School of Life Sciences, Arizona State University, USA

Abstract

Despite the tremendous diversity and complexity of life forms, there are certain forms of life that are never observed. Organisms like angels might not emerge because of developmental biases that restrict how organisms can evolve, or because they have low fitness in any environment yet available on Earth. Given that both developmental bias and selection may create similar phenotypes, it is difficult to distinguish between the two causes of evolutionary stasis among related taxa. For example, remarkably invariant traits are observed spanning million years, such as wing shape in *Drosophila* wherein qualitative differences are rare within genera. We thus ask whether the absence of combinations of traits, indicated by genetic correlation, reflects developmental bias limiting the possibility of change. However, much confusion and controversy remain over definitions of developmental bias and quantifying it is challenging. We present a novel approach aiming to estimate developmental bias by leveraging a common but under-utilized type of data: recombinant genetic mapping populations. We reason that information rendered by such mild perturbations captures inherent interdependencies between traits – developmental bias. Through empirical analyses, we find that our developmental bias metric is a strong indicator of genetic correlation stability across conditions. Our framework presents a feasible way to quantify developmental bias between traits and opens up the possibility to dissect patterns of genetic correlation.

Keywords: Pleiotropy; Genetic correlation; Developmental bias; Evolvability

Significance Statement

Genetic correlation represents an important class of evolutionary constraint, which are themselves evolvable. Empirical studies have found mixed results on whether such evolutionary constraint changes rapidly or slowly. This uncertainty challenges our ability to predict the outcome of selection. Here, we propose a framework to dissect genetic correlation in a genetic mapping population and show that consistency of pleiotropic effects of loci across the genome, which we termed as developmental bias, is an indicator of genetic correlation stability. Our novel method empowers readily accessible QTL mapping data to understand complex genetic architecture underlying pleiotropy, mechanisms causing genetic correlation and, ultimately, long-term evolutionary divergence.

Introduction

> *When the morphology of a species remains virtually unchanged for millions of years, we would like to know whether this reflects developmental constraints limiting the possibility of change or, conversely, the maintenance of uniformity by stabilizing selection.* — *Maynard Smith* et al. *1985*

Genetic correlation represents an important class of evolutionary constraints (Maynard Smith *et al.*, 1985; Clark, 1987), affecting future evolutionary trajectories. Yet, genetic correlations are themselves evolvable (Doroszuk *et al.*, 2008; Dugand *et al.*, 2021; Delph *et al.*, 2011; Conner, 2002; Uller *et al.*, 2018; Wagner & Altenberg, 1996; Rohner & Berger, 2023; Wagner *et al.*, 2007) and reflect both the past selection of trait combinations and, in some cases, developmental bias (Dugand *et al.*, 2021; Arnold, 1992). Natural selection may favor certain combinations of traits and thereby actively maintaining genetic correlation via pleiotropy or linkage disequilibruim. Pleiotropy and linkage disequilbruim (LD) may then further inhibit traits from evolving independently towards a theoretical phenotypic optimum (Schluter, 1996). On the other hand, genetic correlation can be shaped by bias due to intrinsic attributes of the organism, energy, or the laws of physics, relative to the assumption of isotropic variation. This latter concept has been decribed as developmental constraint or developmental bias (Maynard Smith *et al.*, 1985; Arnold, 1992; Cheverud, 1984; Rohner & Berger, 2023), which may account for the observation that perturbations, such as mutation or environmental variation, to biological systems will tend to produce some phenotypic variants more readily (Uller *et al.*, 2018; Waddington, 1957). In spite of the numerous studies that address genetic correlation as an evolutionary constraint, much confusion and controversy remains over definitions of dif-

ferent types of constraint, the mechanism(s) causing constraint, and the relative importance of different mechanisms in shaping evolutionary trajectories (Muir *et al.*, 2022; Conner *et al.*, 2011).

The theoretical underpinnings for genetic covariance as an evolutionary constraint are well-developed (Lande, 1979; Lande & Arnold, 1983). Genetic covariance specifically describes trait covariance due to pleiotropic alleles, where a single locus has effects on two traits, or due to linkage disequilibrium of two loci, each of which affects a single trait but are physically so close that these two traits are strongly associated in populations (Lande, 1980; Lynch *et al.*, 1998; Falconer *et al.*, 1996; Conner *et al.*, 2004). The genetic information summarized by genetic covariance is connected to evolutionary processes in complex ways. For example, evolution toward a phenotypic optimum for two traits may be restricted if selection favors two traits antagonistically but the traits are positively correlated. That is, adaptive evolution can be limited if the joint vector of selection is antagonistic to the trait correlations. In some cases, such evolutionary constraint may persist over long time scales (McGlothlin *et al.*, 2018; Opedal *et al.*, 2023).

Straightforward applications of evolutionary quantitative genetic theory regarding the joint evolution of a pair of traits generally assume an invariant genetic covariance structure (**G** matrix) over the time frame of interest. However, the stability of genetic covariances and how they evolve remain unclear and contentious (Turelli, 1988; Bürger & Lande, 1994; Arnold *et al.*, 2008; Steppan *et al.*, 2002; Milocco & Salazar-Ciudad, 2022; Loeschcke, 1987; Barton & Turelli, 1989). Empirical studies of the evolution of genetic covariance structure have found mixed results on whether genetic covariance changes rapidly or slowly. Some comparisons of

G matrices between natural populations found no evidence of change in **G** (Delahaie *et al.*, 2017; Arnold *et al.*, 2008; Hangartner *et al.*, 2020; Henry & Stinchcombe, 2023), while others have found changes in genetic covariance in only a few generations, across populations, in response to selection, or across environmental conditions (Chakrabarty & Schielzeth, 2020; Milocco & Salazar-Ciudad, 2022; Eroukhmanoff & Svensson, 2011; Walter *et al.*, 2018; Wood & Brodie III, 2015; Henry & Stinchcombe, 2023; Hudson *et al.*, 2022; Scoville *et al.*, 2009; Monroe *et al.*, 2021). We might also predict that the genetic covariance among some suites of traits is stable, while it is unstable for others (Jones *et al.*, 2003). Generally, it is largely unknown what determines the stability of genetic covariance (Wood & Brodie III, 2015) and this uncertainty challenges our ability to predict the outcome of selection.

The persistence of correlational constraint and whether genetic correlation is a good predictor of long-term evolutionary divergence ultimately hinge on our understandings of the underlying mechanism(s) causing genetic correlation (Loeschcke, 1987; Conner *et al.*, 2011, 2004). For example, genetic correlation due to pleiotropy or tight linkage are much more likely to cause evolutionary constraint than those caused by linkage disequilibrium between loosely linked loci (Conner *et al.*, 2011, 2004; Conner, 2002). Correlations due to pleiotropy or tight linkage may persist in the absence of selection, while correlations caused by linkage disequilibrium can be changed quickly by recombination and selection (Conner, 2002; Conner *et al.*, 2004, 2011). We here reason that genetic correlation due to developmental bias is more likely to impose constraint on evolutionary change and may be more persistent than other factors, as developmental bias may arise due to simple principles of physics or chemistry. Insight into the role of developmental bias may reveal why genetic correlations between some traits are

more constant over a long period as compared to other pairs of traits and why, in some cases, genetic constraints can be readily degraded by natural or artificial selection. However, formally discriminating between developmental bias and other mechanisms of genetic correlation is notoriously difficult (Maynard Smith *et al.*, 1985).

Here, we provide an approximate measure of developmental bias by exploiting recombinant genetic perturbations. We define horizontal pleiotropy to describe a locus that has an effect on two traits, where such pleiotropic effect deviates from the effects of the other loci across a genome (Fig. 1 b,d). Conversely, developmental bias between traits describes the observation of consistent pleiotropic effect of loci throughout the genome on a given trait pair (Fig. 1 a,c). We use this consistency of pleiotropic effect throughout the genome to indicate developmental bias, r_D. We reason that if two traits are correlated because of developmental bias, these two traits should be correlated regardless of which specific variant causes the effect.

Our primary goal in the present work is to dissect genetic correlations to understand to what degree they are driven by developmental bias vs. horizontal pleiotropy. We do so by using both numerical simulations and data from a recombinant genetic mapping population. One key outcome is that we identify loci that demonstrate horizontal pleiotropy. While another recent method exists for doing so (Geiler-Samerotte *et al.*, 2020), our method is unique in that it does not require one to study clonal cells and can therefore be applied to a broader range of organisms. An additional goal of our study is to test our proposition that a genetic constraint that arises principally from developmental bias is more persistent than one arising from horizontal pleiotropy. When r_G are driven by numerious small effect size loci, we expect them to be more representative of inherent relationships, as opposed to when they are driven by

individual horizontal pleiotropic loci. In the latter case, any changes or perturbations affecting that specific loci (e.g., various types of environmental perturbations with QTL-by-environment effect, allele frequency changes etc.) may easily disrupt the genetic constraint. We find evidence that, indeed, our estimated developmental bias is an indicator of genetic correlation stability, suggesting that this may allow us to predict change in a genetic correlation over a long-term period. We also show that genetic correlations are likely driven by developmental bias with a highly polygenic architecture. Hence, a genetic correlation with a highly polygenic architecture may be more stable. In sum, we use readily accessible QTL mapping data to understand how genetic architecture influences the portion of a given genetic correlation attributable to developmental bias, to identify loci that act via horizontal pleiotropy, and to make predictions about how genetic correlations will change. These results suggest that this type of common data is under-utilized, and that analyzing recombinant populations with our approach can help to deepen our understandings of genetic correlation.

Results

During a genetic association study, each genetic marker is assigned an odds likelihood ratio along with a effect size for the trait of interest. Instead of identifying statistically significant loci in such conventional genetic association studies, the essential idea, here, is to examine the consistency of pleiotropic effects across genetic backgrounds. We here quantify the developmental bias, r_D, by examining the additive effect of loci for trait pairs throughout the genome. We define a locus with effects that deviate from the overall bivariate trend throughout the genome as a horizontal pleiotropic (HP) locus (Fig. 1b). We diagnose r_D as the consistency

of pleiotropy across genetic backgrounds excluding HP loci. In a mapping population, allele substitutions at each locus represent non-directed (i.e., random) perturbations of varying directions and magnitudes. The additive effect of many loci thus are considered as random perturbations to an organism. We reason that if the effect size of these perturbations on two traits are highly correlated (excluding HP loci), the developmental bias between the two traits is likely to be strong. To better illustrate the framework we propose, the bivariate effect size distributions under two scenarios are shown (Fig. 1a, b). The locus with a major phenotypic effect that deviates from the overall trend of other loci throughout the genome is a horizontal pleiotropic locus. Conversely, the consistency of pleiotropy (i.e., the overall trend of bivariate effect size distribution) is quantified as developmental bias.

Simulation demonstrating relationships between r_D and r_G

To examine how the estimated developmental bias – characterized by the effect size correlation among loci – relates to genetic correlation (r_G), we first simulated two thousand trait pairs for a given simulated population with 500 individual genotypes. For each trait pair, genetic architecture with 226 loci was generated, with additive effect sizes sampled from a multivariate Laplace distribution. The genetic values are obtained by multiplying the genotypes with allelic effect sizes, assuming no epistasis and no linkage disequilibruim. r_G is calculated by correlating the genetic values between two traits following standard protocols (Falconer *et al.*, 1996). We calculated r_D and corresponding r_G for each pair of traits. Notably, r_G is a correlation across a population of individuals while r_D is a correlation across a population of loci in a genome. Therefore, in princple, under a given r_D, the genetic correlation can vary greatly because of the changing allele frequency (Fig. S1).

There exists considerable debate about regimes of allelic effect sizes and their effects on phenotypic evolution: in small steps, via changes of infinitesimally small effect, or in leaps via rare large effect loci (Orr, 2005). Additionally, classic work suggests that different genetic regimes may affect the rate of changes of genetic (co)variance (Barton & Turelli, 1987, 1989; Lande, 1979). Therefore, in addition to examining the effects of HP and LD on the relationship between developmental bias and genetic correlation, we performed these simulations under two genetic regimes, one with high polygenicity which causes low kurtosis in the distribution of effect sizes, and one with low polygenicity which causes high kurtosis in the distribution of effect sizes (Fig. 2).

Assuming no horizontal pleiotropy (HP) and linkage disequilibrium (LD), we expect r_D and r_G to be equal. As expected, without accounting for HP and LD, r_D strongly correlates with r_G regardless of the genetic regimes (Fig. 2b,c). Next, we repeated our simulations under conditions with HP or LD to understand how these forces would affect the correlations (Fig. 2d,e). Under the HP scenario, n randomly selected SNPs ($0 < n < 10$) are forced to have an HP effect, either concordant to or antagonistic with the rest of loci. The genetic correlation r_G is not perfectly correlated with r_D under scenarios with HP, especially when the kurtosis of the effect size distribution is high. In an extreme case, a single large-effect locus can drive the trait correlation despite the low r_D (Fig. S2). Collectively, these observations suggest that r_D and HP loci are two components of r_G, and that even a single large-effect HP locus can drive r_G without overall consistency of pleiotropy throughout the genome.

To understand how LD affects the relationship between r_G and r_D, we also performed simulations using actual recombinant genotypes from a yeast mapping population (Geiler-Samerotte

et al., 2020) (Fig. 2f,g). Similarly to our simulations above, we sampled the effect size for each SNP on each trait from bivariate Laplace distribution with the same γ and then propagated the effect size for a given SNP by "contaminating" its effect size according to the effect sizes of the SNPs in LD with it. (This procedure only accounts for weak linkage, See Supplementary Note 1.) In these simulations, LD appears to affect r_G with a given effect size correlation even for cases in which the genetic architecture is highly polygenic (Fig. 2f). These results imply that LD does not always strengthen r_G; LD could also weaken r_G when, for example the effect of two loci in LD are antagonistic with the overall trend of pleiotropy across the genome. To summarize the numerical simulations, LD, HP, and r_D together shape r_G. In the absence of HP and LD, r_D, we should not expect r_D to be different from r_G. Furthermore, the effects of LD and HP on genetic correlation can become relatively stronger under a more 'Mendelian' genetic architecture with lower polygenicity.

Identifying horizontal pleiotropic loci and delineating developmental bias for yeast morphological traits

We next applied our approach in a yeast morphology dataset, where 374 recombinant strains of yeast cells were imaged for, on average, 800 fixed, stained cells per strain using high-throughput microscopy (Geiler-Samerotte *et al.*, 2020). In total, measurements of 167 morpholgical traits were acquired. The patterns in this large dataset could offer a empirical picture of how HP and LD affect effect size correlations and how our approach can distinguish two mechanisms causing genetic correlation.

As described above, we define developmental bias r_D as the effect size correlation for a subset of

variants where outliers (HP loci) are removed. For example, the effect size distribution (exclude HP loci) for two pair of traits are shown in Fig. 3. The red lines indicate the magnitude of developmental bias (r_D) and the plot on the right is inferred to have a higher developmental bias. To identify outliers (HP loci), we first calculate the correlation by individual-level product (Lea *et al.*, 2019) for each trait pair across each locus. Outliers are then identified as the product falling outside 1.5 times the interquartile range above the upper quartile and below the lower quartile of the distribution. Since LD can also potentially affect the correlation of effect size, we conducted LD pruning to subset the variants to remove loci highly correlated within the population (See Materials and Methods). Therefore, in total, we present the effect size correlation against r_G in three settings: Default (using all genotyped variants), LD corrected, and outlier corrected (i.e., r_D).

Fig. 4a presents the distribution of effect size correlations using all genotyped variants, only LD pruned variants, or outlier-corrected variants (r_D). LD does not exert effect on the patterns of effect size correlation in this dataset (4a, two-sample Kolmogorov-Smirnov test, D = 0.017006, p-value = 0.3879), but horizontal pleiotropic loci, which we identified as outliers, appear to strengthen the effect size correlation (Fig. 4a,b). In principle, an outlier can either weaken or strengthen the effect size correlation. However, our results suggest a bias towards concordant effects between outliers and other loci, given that the distribution under the outlier-corrected setting (r_D) has smaller variance (Fig. 4a) and effect size correlation is generally weaker under the outlier corrected setting (Fig. 4b). Notably, points that deviate more from the unity line ($y = x$) may represent trait pairs which are more strongly affected by horizontal pleiotropy (Fig. 4c).

To further investigate horizontal pleiotropy (HP), we identified those trait pairs significantly affected by the outliers (i.e., yellow dots in Fig. 4b). Outlier loci for these trait pairs are likely indicative of horizontal pleiotropy. Indeed, we confirmed that our method identifies two loci (L15.9 and L13.7, See Table 1) that presented the strongest evidence of horizontal pleiotropy in an earlier study which used stronger genetic correlation than within-line environmental correlation as an indicator (Geiler-Samerotte *et al.*, 2020). Additionally, we found trait pairs with extremely high effect size correlations lacking evidence of horizontal pleiotropy (Fig. 4a, b); an example effect size distribution of a trait pair with exceptionally high effect correlation – possibly reflecting a strong developmental bias between two traits – is shown in Fig. S4. In summary, we show that horizontally pleiotropic loci may indeed affect r_G in the absence of an exceptionally strong developmental bias and that our method can be used to identify horizontal pleiotropic loci and further delineate r_G.

To assess how effect size correlations under three settings relates to genetic correlations r_G, we calculated the effect size correlations between pairwise traits and plotted them against r_G for each trait pair (Fig. S3). Under all settings, we find no cases where trait pairs with no r_G exhibit a strong effect size correlation, as expected. Qualitatively similar results are observed for an additional *Brassica* dataset with 11 floral, vegetative, and phenology traits (Supplementary Note 2, Fig. S7). Notably, without LD correction and without removing horizontal pleiotropic loci as outliers (Fig. S3a), the results here from recombinant perturbations are algined with a study using mutational accumulation lines to examine the contribution of mutation to genetic correlation, in which mutational correlations between traits were found to be overall stronger than genetic correlations between traits (Dugand *et al.*, 2021). Similarly, in our results, effect

size correlations using all variants overall are stronger than genetic correlations (Wilcoxon signed-rank test, p-value $< 2.2e\text{-}16$), with a median absolute value of 0.280 and 0.206, respectively. After removing horizontal pleiotropic loci as outliers, the dots in scatter plot (Fig. S3c) appears to be more evenly distributed around the unity line. Indeed, after removing HP loci, there is no significant difference between the distribution of effect size correlations (r_D) and the genetic correlation across trait pairs (Wilcoxon signed-rank test, p-value = 0.8875).

Despite that there is no overall difference between the distribution of r_D and r_G, our approach can delineate trait-trait specific mechanims causing their genetic correlations in empirical datasets. As shown in Fig. 3, two pairs of traits exhibit a similar and moderately high r_G but contrasting levels of r_D, demonstrating how consistency of pleiotropy and estimated developmental bias r_D could help us learn the underlying trait-trait specific mechanisms.

r_D predicts the stability of r_G following environmental perturbations

Genetic correlations between traits may alter the evolutionary trajectory of either trait (Schluter, 1996). Predicting the trajectory of trait evolution therefore can depend upon the stability of genetic correlations (Jones *et al.*, 2003). We reasoned that trait-trait correlations may be more stable if they are caused by inherent relationships between the traits, r_D, rather than horizontal pleiotropy. Thus we expected r_D to predict the stability of r_G (Fig. 5a). To test whether our intuition is correct, we estimated r_G from a related yeast dataset, which describes correlations across yeast single-cell morphological features measured in three environments. Here, the environmental conditions are three concentrations of geldanamycin (GdA), a small-molecule inhibitor that binds the ATP-binding site of the chaperone Hsp90, thus rendering

it unable to perform its cellular function. We plotted absolute r_D with changes of genetic correlation (r_G) for each pair of traits at the three drug concentrations (Fig. 5). The results show that, as r_D becomes greater, the changes of r_G become smaller.

Since r_G is highly correlated with r_D, to formally test whether under a given r_G, r_D is informative in determining changes of r_G upon environmental perturbations, we conducted multivariable linear regression ($\Delta r_G \sim r_G + r_D$, all variables are transformed to absolute value). The regression results (Fig. S5a) demonstrate that conditioning on r_G of a trait pair, r_D significantly negatively correlates with the changes of r_G across three drug concentrations. In other words, given a set of yeast morphology trait pairs with the same levels of r_G, the changes of magnitude of r_G would be smaller for trait pairs with larger r_D, on average. Furthermore, we found that this effect of r_D is strongest under mild treatment perturbation (here, low concentration of geldanamycin) but becomes weaker as the drug concentration becomes higher and, presumably, more stressful for the cells (Geiler-Samerotte *et al.*, 2016). To account for the effect of collinearity between r_D and r_G (PCC = 0.939) on regression outcomes, we also report, here, null model simulated results (Fig. S5b and Fig. S6). Taken together, these results indicate that the estimated r_D may indeed predict the stability of r_G following environmental perturbtions.

Discussions

It has long been recognized that developmental integration is one cause of multivariate genetic constraint (Klingenberg, 2005; Pigliucci & Preston, 2004). On the other hand, genetic constraint can also reflect correlational selection. However, dissecting the underlying mecha-

nism(s) causing genetic correlation is challenging. Here, we exploited a hidden source of data that has been overlooked to quantify the contribution of developmental bias in creating genetic correlation. Assessing consistency of pleiotropy by measuring the effect size correlation across many genomic loci provides a possible framework to explore the mechanisms of genetic correlation. The central messages from our analyses are three-fold.

First, developmental bias estimated from recombinant genetic perturbations provides an indicator of genetic correlation stability following environmental perturbations. Our stability analyses in empirical datasets (Fig. 5, Fig. S5, Fig. S6, Fig. S8, and Fig. S9) suggest that the higher a developmental bias, the more likely a genetic correlation between two traits remains stable across environmental conditions. In other words, higher developmental bias leads to smaller response of genetic correlation to environmental changes. This may provide further insight into the observations of context-dependencies of environmental effect on **G**-matrices (Wood & Brodie III, 2015), with certain trait pairs exhibiting more stability while others showing greater plasticity across conditions.

Second, Mendelian genetic architecture for a given trait pair can increase the contribution of horizontal pleiotropy to genetic correlation (Fig. 2). Under such a scenario, genetic correlation as a summary statistic can not fully reflect the complex genetic architecture underlying a genetic correlation. In fact, evidence of discrepancies of effect between genetic background and major loci abound (Albert *et al.*, 2008; Hall *et al.*, 2006; Scoville *et al.*, 2009; Stinchcombe *et al.*, 2009). For example, in *Mimulus*, a major QTL contributes a negative covariance between stigma–anther separation and pollen viability, which is antagonistic to the overall positive genetic covariance between these two traits (Scoville *et al.*, 2009). Furthermore, previous work

suggests that we might expect to see more changes of **G**-matrix during evolution if traits have an oligogenic genetic basis rather than aligning with the infinitesimal model (Barton & Turelli, 1987, 1989; Lande, 1979). For example, Lande (Lande, 1979) emphasized that trait means typically change much more rapidly than trait (co)variances. Yet, changes of (co)variances can be quite rapid if there are underlying loci with large contributions to (co)variation. Similarly, in our present work, we show that under a infinitesimal model, genetic correlation mainly arises from developmental bias (Fig. 2) which, as our stability tests suggest, might also be more stable across conditions (Fig. 5).

Third, our method allows us to identify horizontal pleiotropic loci without measuring phenotypes across clonal individuals or cells. There is a long-standing interest in identifying horizontal pleiotropy in nature (Verbanck *et al.*, 2018; Jordan *et al.*, 2019; Bowden *et al.*, 2018). One motivation for doing so is that evolutionary theory predicts that natural selection should limit horizontal pleiotropy because, as the number of traits that a mutation influences increases, the probability of the mutation having a positive fitness effect decreases (Zhang & Wagner, 2013; Orr, 2000; Pavlicev & Wagner, 2012; McGuigan *et al.*, 2014). However, identifying cases of horizontal pleiotropy is difficult because genetic correlations do not always indicate horizontal pleiotropy. By discovering a way to disentangle the portion of genetic correlation caused by developmental bias, we have also discovered a novel way to identify candidate loci that act via horizontal pleiotropy. Our method of identifying horizontal pleiotropy can be broadly useful because it does not require measuring the trait correlations that are present across clonal cells. Thus, while previous methods (Geiler-Samerotte *et al.*, 2020) are mainly useful for organisms that propagated clonally, e.g., microbes, our method can be applied more

broadly.

Can recombinant mapping population characterize M matrix?

Numerous past studies used mutation accumulation lines to estimate mutational matrices (**M** matrices) as a means to understand the influence of mutations on shaping genetic correlation (Dugand *et al.*, 2021; Houle *et al.*, 2017). Dugand *et al.* (2021) discovered a significant similarity between **G** and **M** matrix, suggesting that mutations directly shape **G**. On the other hand, mutational correlations using mutation accumulation lines consistently exceed genetic correlations in magnitude (Dugand *et al.*, 2021), which is aligned with our findings (Fig. 3c) where the effect size correlations under default setting are stronger than the genetic correlation r_G. This naturally raises several questions related to mutation accumulation lines, recombinant mutation, and developmental bias: Firstly, to what extent does recombinant reflect the effect of mutation in a mutational accumulation experiment? There are now increasingly accessible resources available for recombinant mapping populations, such as the recently developed multiparent panels and advanced intercross lines (Kover *et al.*, 2009; Gage *et al.*, 2020), offering a promising avenue to investigate the respective roles of mutation and selection. Secondly, to what extent do trait covariance patterns due to mutations or environmental perturbations reflect the developmental bias? A recent study used fluctuating asymmetry of the left and right sides of the same organism as a measure of developmental bias (Rohner & Berger, 2023). The left and right sides of the same organism share the same genome and macro-environment but only differ in their microenvironmental inputs. Therefore, the development may generate asymmetry (i.e., noise) in morphological traits. The authors showed that developmental bias quantified using such noise in the dipteran wing predicts its evolution on both short and long

evolutionary timescales(Rohner & Berger, 2023), which suggests that those mild perturbations may generate phenotypic outcomes more representative of developmental bias.

The extent of pleiotropy

Pleiotropy describes the phenomenon in which a gene or a mutation affects more than one phenotypic trait. The concept and nuance of pleiotropy has had a prominent role and broad implications on genetics, evolution, and medicine (Klingenberg, 2008; Stearns, 2010; Promislow, 2004; Williams, 2001; Barton, 1990; He & Zhang, 2006; Otto, 2004; Wagner & Zhang, 2011; Des Marais & Juenger, 2010; Geiler-Samerotte *et al.*, 2020). Conceptually, many possible scenarios can result in a pleiotropic effect, including mediated pleiotropy (i.e., vertical pleiotropy), horizontal pleiotropy, and other spurious pleiotropy such as linkage (Wagner & Zhang, 2011; Solovieff *et al.*, 2013). One major debate on pleiotropy is what is the extent of pleiotropy: we lack consensus about how pleiotropic natural systems are (Paaby & Rockman, 2013; Zhang & Wagner, 2013). A key challenge is whether the effect of a single locus on correlated traits can be counted as pleiotropic effect, for instance, as pointed out by Wagner & Zhang 2011; e.g., are the depth and the width of a bird beak two characters? Thus, ignoring trait correlations may bias the estimation of pleiotropy. One possible solution is to consider the effective number of traits by looking at the eigenvalue variance of the phenotypic correlation matrix (Wagner & Zhang, 2011; Pavlicev *et al.*, 2009; Wagner *et al.*, 2008): The more dispersed the eigenvalues, the more interdependency of the traits. However, this approach likely biases the interdependency estimations of the traits, especially in the presence of major pleiotropic effect loci. For example, as an extreme case, even a single pleiotropic locus alone can drive trait correlations in spite of the low consistency of pleiotropy (Fig. S2 and also see (Agrawal

et al., 2001)). Hence, in this case, the effective number of traits calculated via the phenotypic correlation matrix will be overestimated simply because there is a major effective pleiotropic locus – this does not necessarily mean two traits are inherently interrelated. Similarly, the bias is present if there is a major antagonistic loci against overall correlation of effect size. Instead, our analyses demonstrated that the consistency of effect sizes may provide a more appropriate way to measure inherent trait correlation and hence effective trait dimensions.

Materials and Methods

Sophisticated tools in the field of quantitative genetics have been developed to identify genetic loci which statistically explain phenotypic variance in quantitative traits to regions of chromosomes, so-called quantitative trait loci (QTLs). One of the fundamental metrics of quantitative genetics is the additive effect of a QTL, which represents the change in the average phenotype produced by substituting one allele for another (Lynch *et al.*, 1998; Falconer *et al.*, 1996). To better illustrate what we could exploit through the additive effect distribution, bivariate effect size distribution under two scenarios are shown (Fig. 1a,b), where both of two pairs of traits are affected by a major pleiotropic locus. In contrast, the consistency of pleiotropic effect throughout genome is different. This illustrative example may be extreme, but it implies that only analyzing the summary statistics such as genetic correlation or statistically significant loci in a genetic association study may lose information behind the genetic architecture. Such hidden information could be valuable when assessing the strength of developmental bias r_D: If two traits are correlated because of developmental or physiological constraint, these two traits should be correlated regardless of which specific variant is causing the effect. i.e., there

is consistency of the pleiotropic effect across genetic background (Fig. 1a). On the other hand, if two traits are genetically correlated simply because of several major pleiotropic loci for a given population, those small loci can have inconsistent effect between traits (Fig. 1b). We term such consistency of pleiotropy as developmental bias r_D and those loci with effect deviated from overall trend throughout the genome as horizontal pleiotropic (HP) loci.

Our conceptualization of developmental bias is similar to the definition of vertical pleiotropy or mediated pleiotropy Geiler-Samerotte *et al.* (2020). Indeed, the high consistency of pleiotropic effect implies vertical or mediated pleiotropic nature of loci. Yet, we here define the developmental bias as a trait-level metric, whereas the vertical or mediated pleiotropy most often refers to the effects of variants on traits. In vertical pleiotropy, the traits themselves are biologically related, such that a variant's effect on trait A inevitably causes the effect on trait B. Likewise, horizontal pleiotropy is defined as a variant or mutation causing an effect on two traits that are otherwise independent. Another distinction between developmental bias and vertical pleiotropy is that vertical pleiotropy frequently refers to a part of causal cascade, as exemplified by low-density lipoprotein (LDL) cholesterol levels causing the risk of heart disease (Geiler-Samerotte *et al.*, 2020). Developmental bias, on the other hand, depicts the correlational structure among traits since many traits (e.g., morphological traits) do not necessarily exhibit direct causal relationships. We thus apply the term vertical pleiotropy to variants that share the effect for inherently related traits without considering the causal direction.

Numerical simulations

To simulate multiple pairs of traits within a population, first genotypes were simulated through the function *simulateGenotypes* in PhenotypeSimulator (Meyer & Birney, 2018) with 226 SNPs (mimicking the actual number of loci in an empirical dataset; Geiler-Samerotte *et al.* (2020)) and 500 individuals, where the allele frequencies are either sampled from 0.05, 0.1, 0.2, and 0.5 or constant value 0.3 for a mapping population. For each pair of traits, the additive effect for each SNP is sampled from a bivariate exponential distribution (bivariate Laplace distribution) with $\mu = (0, 0)$ and $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. ρ is drawn from the uniform distribution $(-1, 1)$. A shape parameter γ determines the distribution, where a smaller γ represents genetic architecture approximating one or a small number of large-effect loci (Mendelian genetic architecture, high kurtosis for effect size distributions) while a larger γ trends towards a polygenic infinitesimal model (low kurtosis for effect size distributions). We used γ of 1.0 and 0.5 in Fig. 2 left and right, respectively. Under horizontal pleiotropy scenario, n randomly selected SNPs ($0 < n < 10$) are forced to have horizontal pleiotropic effect (either concordant to or antagonistic with the rest of loci). The genetic correlation, r_G, between traits M and N was calculated as the Pearson correlation $\rho_{(X\beta_{[M]}^{\intercal}, X\beta_{[N]}^{\intercal})}$, where $\beta_{[M]}$ and $\beta_{[N]}$ represents the effect size for trait M and trait N across genome, and $X\beta_{[M]}^{\intercal}$ and $X\beta_{[N]}^{\intercal}$ represent the genotypic values of trait M and trait N, respectively. The developmental bias, r_D, is calculated as the Pearson correlation coeffecient of effect size for traits M and N, where summations are taken over all loci except those assigned as horizontal pleiotropic SNPs (n loci):

$$D_{\beta[M,N]} = \frac{cov(\beta_{[M]}, \beta_{[N]})}{\sigma_{\beta_{[M]}}\sigma_{\beta_{[N]}}} = \frac{n\sum\beta_{i[M]}\beta_{i[N]} - \sum\beta_{i[M]}\sum\beta_{i[N]}}{\sqrt{n\sum\beta_{i[M]}^2 - (\sum\beta_{i[M]})^2}\sqrt{n\sum\beta_{i[N]}^2 - (\sum\beta_{i[N]})^2}} \quad (1)$$

To account for linkage disequilibrium (LD), simulating trait pairs with actual genotype information, the additive effect size of 226 SNPs are simulate similarly as above except that there are no horizontal pleiotropic loci. The effect size is then propagated through the LD block defined by $r > 0.5$.

Dataset retrieval and genetic correlations r_G

Two empirical datasets were used. The first dataset comprises single cell morphology data for budding yeast *Saccharomyces cerevisiae* where, for each of 374 recombinant strains of yeast cells, approximately 800 fixed, stained cells were imaged using high-throughput microscopy (Geiler-Samerotte *et al.*, 2020). 167 morphological features were estimated, including these representative examples: cell size, bud size, bud angle. Analysis of the original dataset assessed both genetic (between-strain) and environment (within-strain) correlation using a multilevel correlation partitioning method (Bliese, 2013). The authors found that using this approach to estimate correlations has similar results as compared to a linear mixed model and variance component analysis. The second dataset contains phenological, floral, and vegetative traits for a recombinant inbred population of *B. rapa L.* created from a cross between yellow sarson, R500, and the rapid cycling IMB211 inbred lines (Brock *et al.*, 2010). The QTL mapping was conducted with 223 markers in 131 individuals (field condition) and 132 individuals (greenhouse condition). Eleven phenotypes were included, here, as we excluded branch length in the field and leaf width in the greenhouse, which were not measured in both conditions. When

calculating r_D, we use either all genome-wide variants or an LD-pruned subset of variants. During LD pruning, we removed within-choromosome QTLs to $r < 0.5$ for both datasets.

Horizontal pleiotropic loci identification and empirical calculation of r_D

If we ignore dominance, epistasis, and linkage disequilibrium, and assume two alleles per locus, the covariance components of a **G** matrix can be written as (Kelly, 2009):

$$C_{\alpha[M,N]} = \sum_i 2q_i(1-q_i)\alpha_{i[M]}\alpha_{i[N]} \tag{2}$$

where $C_{\alpha[M,N]}$ is the additive genetic covariance between trait M and trait N. q_i is the frequency of first allele at loci i within a given population; $\alpha_{i[M]}$ and $\alpha_{i[N]}$ are the additive effects of that allele on trait M and N, respectively; summations are taken over all loci. Accordingly, a large effect QTL for trait M (high $\alpha_{i[M]}$) can make a minor contribution to the genetic covariance structure if allele frequency q_i is small.

We developed an approach to evaluate the horizontal pleiotropy and calculate r_D empirically. In brief, our method has three components: (a) detection of horizontal pleiotropic loci; (b) calculating r_D through effect size correlation excluding horitonal pleiotropy; (c) testing pairs with significant difference after outlier removal, identified as horizontal trait pairs.

We use the following procedures to identify horizontal pleiotropic loci: first calculate the normalized, demeaned, and element-wise product of outcome for each locus (Steiger, 1980):

$$\frac{(\alpha_{i[M]} - \overline{\alpha_{[M]}})(\alpha_{i[N]} - \overline{\alpha_{[N]}})}{\sigma_{\alpha_{i[M]}}\sigma_{\alpha_{i[N]}}} \tag{3}$$

where $\overline{\alpha_{[M]}}$ and $\overline{\alpha_{[N]}}$ are the mean genome-wide additive effect size for trait M and N, respectively, and $\sigma_{\alpha_{i[M]}}$ and $\sigma_{\alpha_{i[N]}}$ are standard deviations of additive effect size for trait M and N, respectively. The horizontal pleiotropic loci are defined as loci with the product falling outside 1.5 times the interquartile range above the upper quartile and below the lower quartile. The Pearson correlation coefficient is equal to the above average element-wise product of two measured traits. In other words, the outliers of these element-wise products represent the outliers when calculating the correlation, substantially deviating from the overall trend of bivariate effect size distribution.

Second, developmental bias (r_D) is calculated as the Pearson correlation coefficient among the rest of loci written by:

$$r_{D_{\alpha[M,N]}} = \frac{cov(\alpha_{[M]}, \alpha_{[N]})}{\sigma_{\alpha_{[M]}}\sigma_{\alpha_{[N]}}} = \frac{n\sum_i \alpha_{i[M]}\alpha_{i[N]} - \sum \alpha_{i[M]} \sum \alpha_{i[N]}}{\sqrt{n\sum \alpha_{i[M]}^2 - (\sum \alpha_{i[M]})^2}\sqrt{n\sum \alpha_{i[N]}^2 - (\sum \alpha_{i[N]})^2}} \tag{4}$$

The additive effect size α estimated empirically from inbred line crosses and experimental mapping populations is calculated as: for the trait M, $\alpha = \frac{\overline{M_{AA}} - \overline{M_{BB}}}{2}$ (Falconer *et al.*, 1996), where A and B are two alleles of a locus. The summations of equation (4) are taken across all loci (with equal probabilities) excluding major horizontal pleiotropic loci.

Finally, to test the significant difference of effect size correlation before and after outlier removal, we use a cutoff of 1% FDR for all pairs of traits included in the yeast dataset. Those

trait pairs significantly deviating from the mean are identified as the horizontal trait pair, indicating that horizontal pleiotropic loci may contribute to the genetic correlation r_G.

Data Availability

Data and code have been deposited in Github (`https://github.com/haorancai/developmentalbias`)

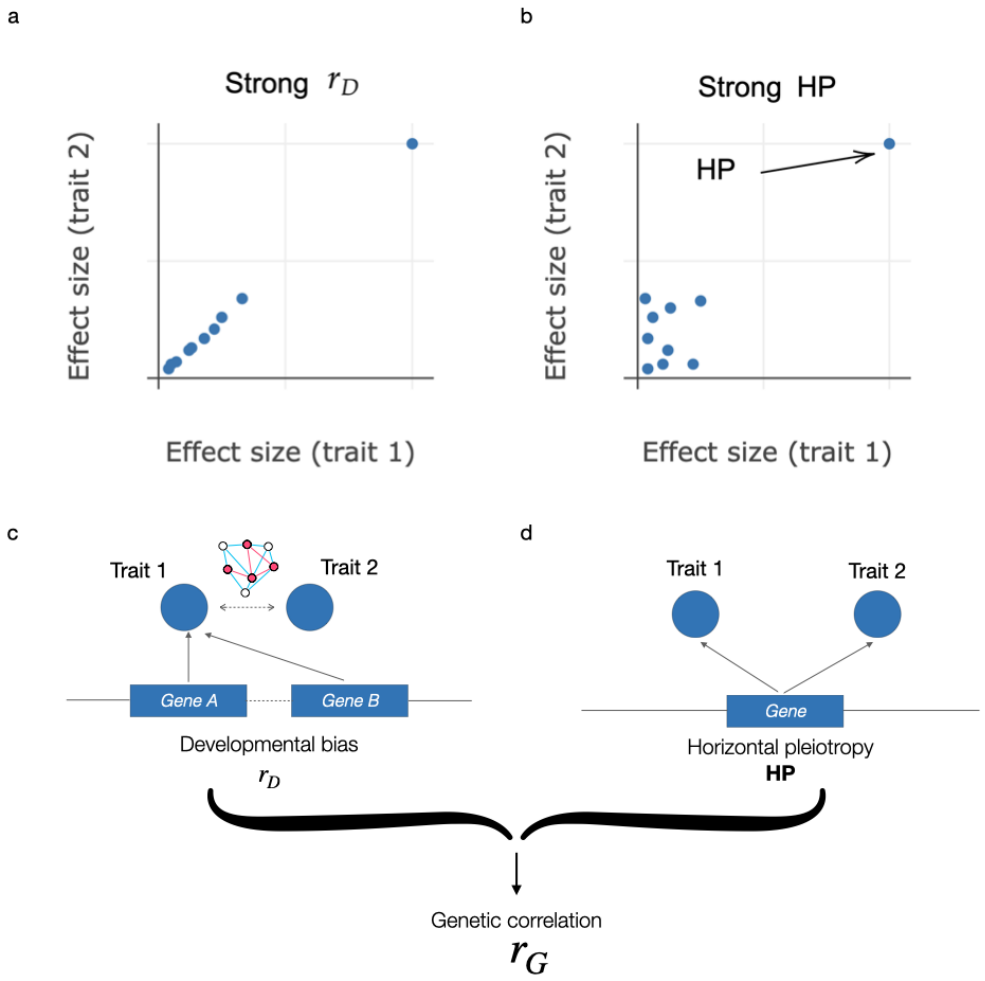

Figure 1: Conceptual framework for distinguishing between developmental bias and horizontal pleiotropy as drivers of genetic correlation (r_G) between two traits. **a.** and **b.** Bi-plots showing the correlation of effect sizes of ten genetic loci on hypothetical traits 1 and 2. **a.** Strong developmental bias and low horizontal pleiotropy, as seen by coherent and consistent pleiotropic effect on the two traits across the sampled loci. **b.** One large-effect pleiotropic locus appears to drive the genetic correlation between traits 1 and 2, showing a strong horizontal pleiotropic (HP) effect. **c.** and **d.** Suggest genetic mechanisms for the observed effect correlations in **a** and **b**. **c.** A developmental bias r_D, where each locus that affects one trait will inevitably affect the other trait, suggesting that the traits are inherently correlated regardless of the type and directions of genetic perturbation. **d.** Horizontal pleiotropy (HP), where a locus can have a direct effect on the two traits. A third cause of genetic correlation is linkage disequilibrium (LD; not shown), where two loci, each determining a trait, are physically so closed that they are associated within a population more often.

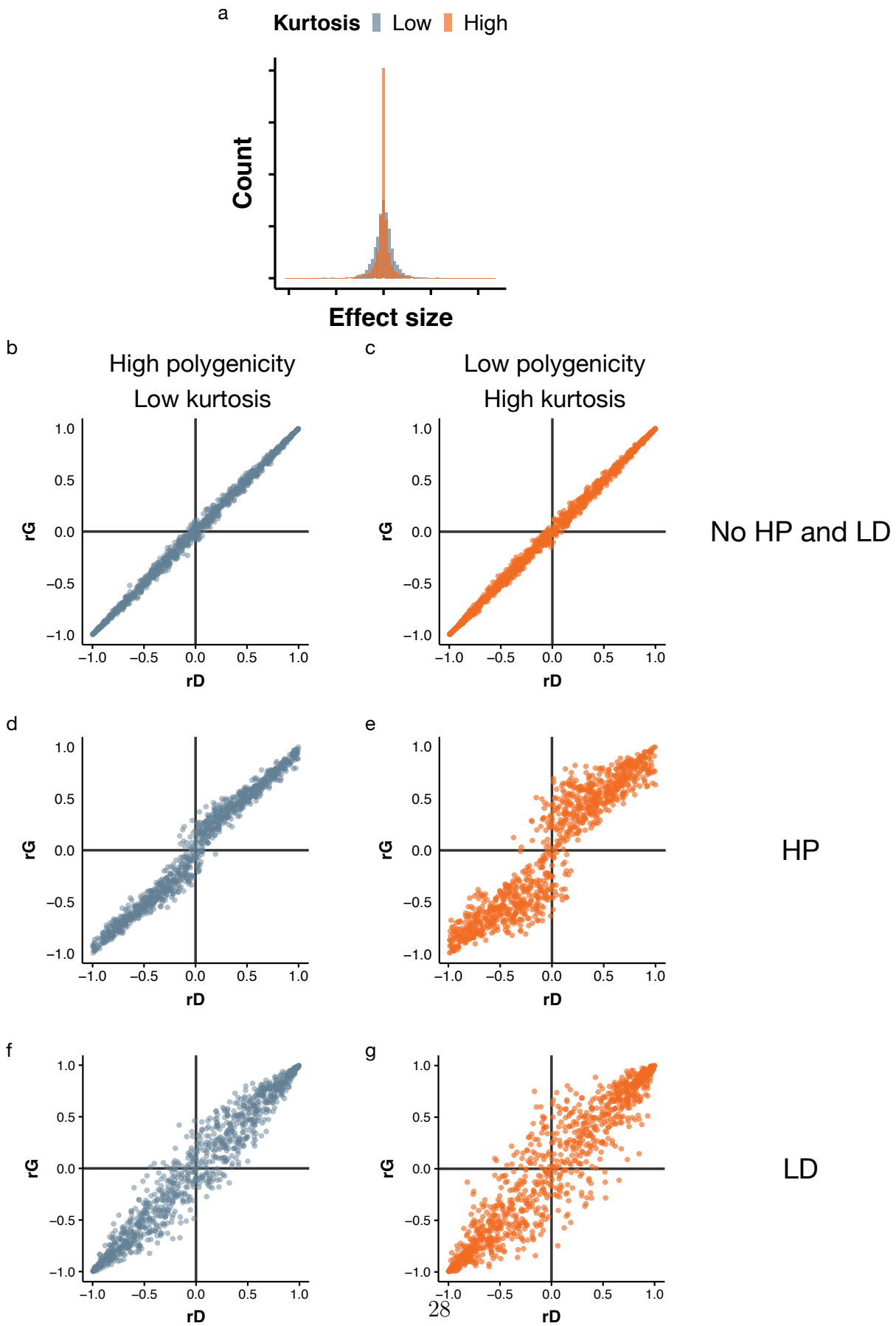
a
Kurtosis
Low
High
Count
Effect size
b
High polygenicity
Low kurtosis
c
Low polygenicity
High kurtosis
No HP and LD
d
e
HP
f
g
LD
rG
rD
1.0
0.5
0.0
−0.5
−1.0

Figure 2: Simulations showing the relationship between developmental bias, r_D, and genetic correlation, r_G. We simulated 2,000 pairs of traits using an exponential model, with varying σ among the plots shown. σ describes the correlation of effect size when generating effect sizes for each trait pair. A second parameter of the exponential model, γ, indicates regimes of varying kurtosis of the effect size distribution: large γ represents low kurtosis while smaller γ represents a regime with high kurtosis, as shown in **a.**. Two regimes of genetic architecture are considered, with γ equals 1.0 (plots **b.**, **d.**, and **f.**) and 0.5 (plots **c.**, **e.**, and **g.**) Three scenarios are simulated. **b.** and **c.** Under the first scenario, no horizontal pleiotropy (HP) and linkage disequilibrium (LD), the pattern does not change with the kurtosis and r_G nearly perfectly represents r_D. **d.** and **e.** Under the second scenario, to introduce HP, 0 to 10 randomly chosen SNP are introduced for each trait pair, with a shared pleiotropic effect for two traits, regardless of the pleiotropic effect of remaining loci. **f.** and **g.** For the third scenario, we used actual genotypes from a yeast mapping population to account for LD, without introducing any horizontal pleiotropic loci. To simulate trait pairs with actual genotype information, effect sizes are first generated, as above, but are propagated through the LD block defined by $r > 0.5$. Note that the correlation of effect size is then calculated by excluding those 'repeated' loci within LD blocks.

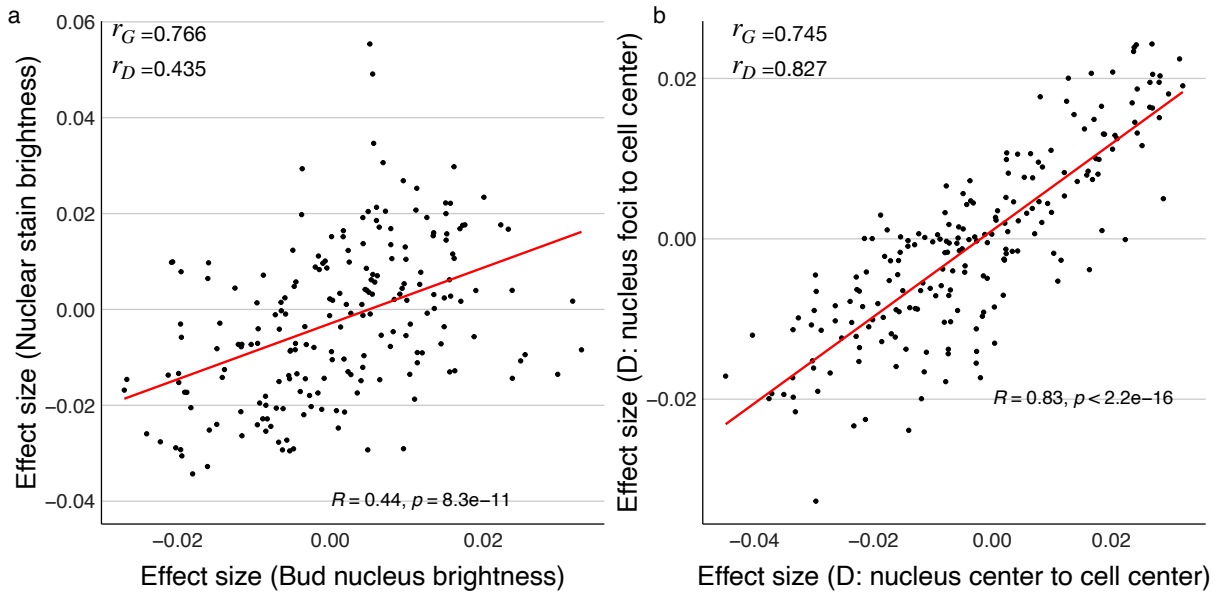

Figure 3: Empirical examples under two scenarios in Fig. 1 demonstrating how estimated r_D can differentiate two trait pairs with similar r_G in yeast. These two trait pairs with similar strength of r_G (0.766 and 0.745) exhibit difference in r_D (0.435 and 0.827). **a.** The r_G between nuclear stain brightness and bud nucleus brightness is marginally higher than the **b.** r_G between nucleus foci-to-cell center and nucleus center-to-cell center. However, r_D between nuclear stain brightness and bud nucleus brightness is lower than r_D between nucleus foci to cell center and nucleus center to cell center (0.435 vs 0.827), suggesting a higher developmental bias and inherent correlation between nucleus center-to-cell center and nucleus foci-to-cell center. Each point represents the additive effects for a single locus on each of two traits shown. Data are from (Geiler-Samerotte *et al.*, 2020)

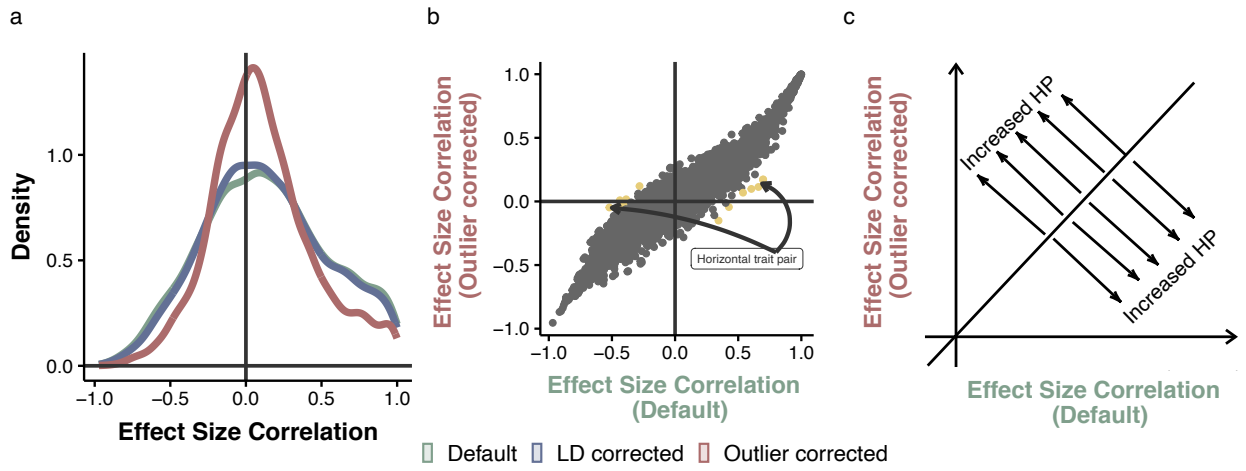

Figure 4: Re-analysis of 374 recombinant strains of yeast cells (Geiler-Samerotte *et al.*, 2020) identifies horizontal pleiotropic trait pair. Each point in **b.** represents a trait pair from this empirical dataset. We consider three settings, summarized in **a.** and **b.**. **a.** Distribution of effect size correlations under three settings. Under the default setting, we included genome-wide markers to calculate correlations of effect sizes. LD pruned results include only the loci to the $r < 0.5$ within a chromosome. For outlier corrected effect size correlation, we excluded those outlier horizontal loci when calculating the correlation coefficient. **b.** The effect size correlation under default versus outlier corrected settings. Yellow dots denote trait pairs that are significantly affected by outliers correction (p-value <0.025). **c.** Conceptual figure showing where in the scatterplot **b** trait-trait effect size correlations have more contribution from horizontal pleiotropy.

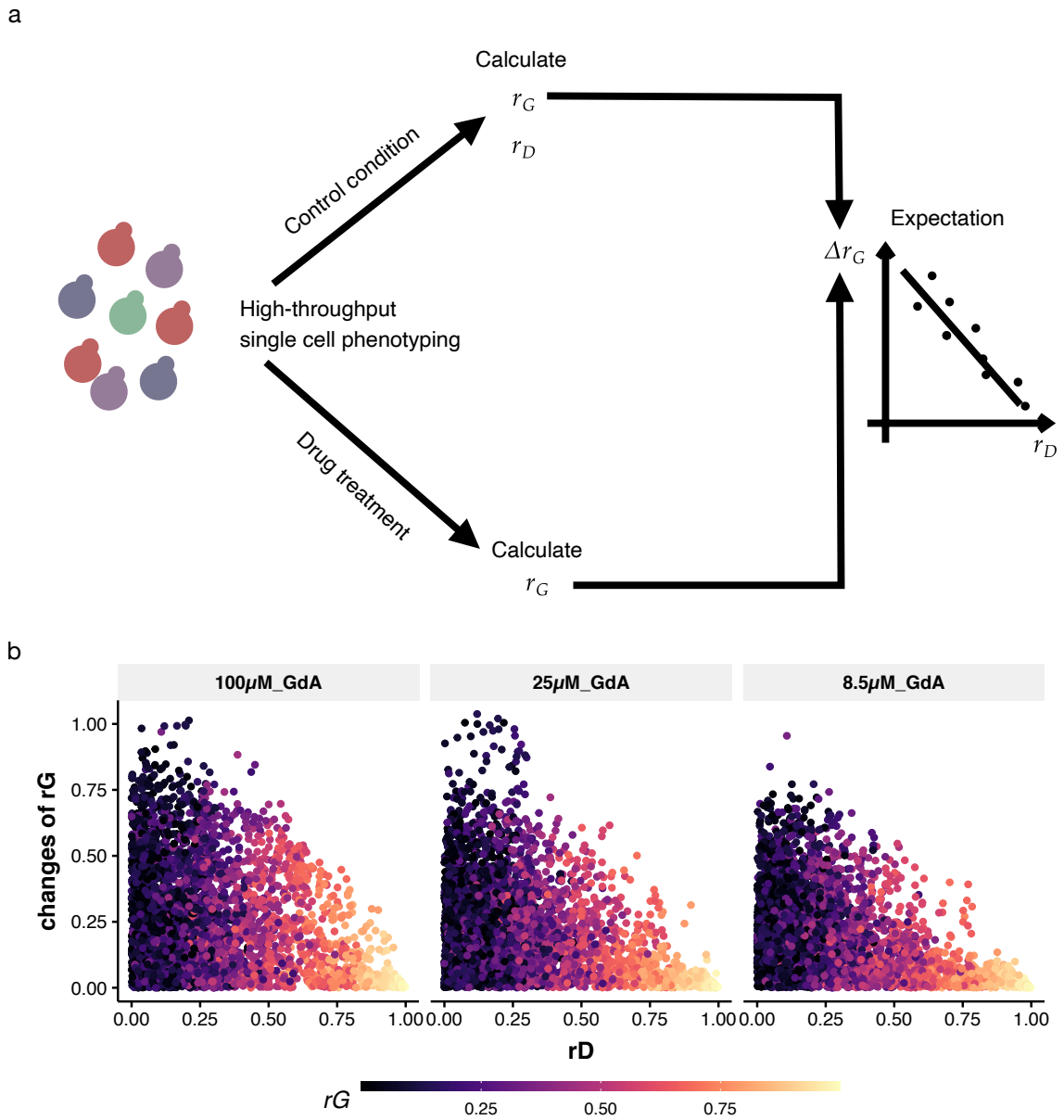

Figure 5: **a** Conceptual representation showing how high r_D could lead to stability of r_G. Many morphological traits are measured in a yeast biparental population using single cell phenotyping (Geiler-Samerotte *et al.*, 2020). These data are used to estimate r_G and r_D between traits. A subset of the yeast population is subjected to three levels of drug concentrations, representing three environmental conditions, and r_G is calculated for traits expressed in each of these treatments. We ask whether r_D is an indicator of Δr_G by using a multivariable linear model: $\Delta r_G \sim r_D + r_G$. We include r_G as a predictor because r_G itself can reflect its own plasticity. **b.** Scatter plots showing how r_G plasiticity in response to different drug concentration relates to r_D. The color bar indicates r_G in control condition.

Table 1: Details of horizontal trait pairs and nearest markers of driver loci

Phenotype1	Phenotype2	rG	rD_outlier_corrected	rD_default	Nearest marker for the biggest outlier
D14.3_A1B	D15.3_A1B	0.458	0.127	0.663	L15.9
D15.3_A1B	D175_A1B	0.447	0.193	0.701	L15.9
D15.3_A1B	D178_A1B	0.448	0.166	0.702	L15.9
D15.3_A1B	D181_A1B	0.372	0.087	0.607	L15.9
C13_C	C103_C	-0.332	0.017	-0.388	L13.7
C13_C	C118_C	-0.087	0.092	-0.29	L13.7
C13_C	D109_C	-0.215	0.017	-0.44	L13.7
C13_C	D131_C	-0.23	-0.055	-0.442	L13.7
D15.1_C	D17.1_C	0.188	-0.052	0.431	L15.9
D15.1_C	D17.2_C	-0.038	-0.023	0.407	L15.9
D15.1_C	D182_C	0.218	0.097	0.529	L15.9
D15.3_C	D17.1_C	0.106	-0.15	0.352	L15.9
D15.3_C	D188_C	0.111	0.096	0.542	L15.9
D17.1_C	D117_C	-0.379	-0.049	-0.523	L15.9
D17.1_C	D169_C	-0.307	0.035	-0.408	L15.9

References

Agrawal, A. F., Brodie, E. D. & Rieseberg, L. H. (2001). Possible consequences of genes of major effect: transient changes in the g-matrix. In: *Microevolution rate, pattern, process*. Springer, pp. 33–43.

Albert, A. Y., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., Balabhadra, S., Kingsley, D. M. & Schluter, D. (2008). The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution: International Journal of Organic Evolution*, 62, 76–85.

Arnold, S. J. (1992). Constraints on phenotypic evolution. *The American Naturalist*, 140, S85–S107.

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

Ashman, T. & Majetic, C. (2006). Genetic constraints on floral evolution: a review and evaluation of patterns. *Heredity*, 96, 343–352.

Ashman, T. L. (1999). Quantitative genetics of floral traits in a gynodioecious wild strawberry fragaria virginiana: implications for the independent evolution of female and hermaphrodite floral phenotypes. *Heredity*, 83, 733–741.

Barton, N. (1990). Pleiotropic models of quantitative variation. *Genetics*, 124, 773–782.

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

Barton, N. H. & Turelli, M. (1989). Evolutionary quantitative genetics: how little do we know. *Annual review of genetics*, 23, 337–370.

Bliese, P. (2013). Multilevel modeling in r (2.6). *Retrieved September*, 3, 2013.

Bowden, J., Hemani, G. & Davey Smith, G. (2018). Invited commentary: detecting individual and global horizontal pleiotropy in mendelian randomization—a job for the humble heterogeneity statistic? *American journal of epidemiology*, 187, 2681–2685.

Brock, M. T., Dechaine, J. M., Iniguez-Luy, F. L., Maloof, J. N., Stinchcombe, J. R. & Weinig, C. (2010). Floral genetic architecture: an examination of qtl architecture underlying floral (co) variation across environments. *Genetics*, 186, 1451–1465.

Bürger, R. & Lande, R. (1994). On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics*, 138, 901–912.

Chakrabarty, A. & Schielzeth, H. (2020). Comparative analysis of the multivariate genetic architecture of morphological traits in three species of gomphocerine grasshoppers. *Heredity*, 124, 367–382.

Chebib, J. & Guillaume, F. (2021). Pleiotropy or linkage? their relative contributions to the

genetic correlation of quantitative traits and detection by multitrait gwa studies. *Genetics*, 219, iyab159.

Cheverud, J. M. (1984). Quantitative genetics and developmental constraints on evolution by selection. *Journal of theoretical biology*, 110, 155–171.

Clark, A. (1987). Genetic correlations: the quantitative genetics of evolutionary constraints. In: *Genetic constraints on adaptive evolution*. Springer.

Conner, J. K. (2002). Genetic mechanisms of floral trait correlations in a natural population. *Nature*, 420, 407–410.

Conner, J. K., Hartl, D. L. *et al.* (2004). *A primer of ecological genetics*, vol. 425. Sinauer Associates Sunderland, MA.

Conner, J. K., Karoly, K., Stewart, C., Koelling, V. A., Sahli, H. F. & Shaw, F. H. (2011). Rapid independent trait evolution despite a strong pleiotropic genetic correlation. *The American Naturalist*, 178, 429–441.

Delahaie, B., Charmantier, A., Chantepie, S., Garant, D., Porlier, M. & Teplitsky, C. (2017). Conserved g-matrices of morphological and life-history traits among continental and island blue tit populations. *Heredity*, 119, 76–87.

Delph, L. F., Steven, J. C., Anderson, I. A., Herlihy, C. R. & Brodie III, E. D. (2011). Elimination of a genetic correlation between the sexes via artificial correlational selection. *Evolution*, 65, 2872–2880.

Des Marais, D. L. & Juenger, T. E. (2010). Pleiotropy, plasticity, and the evolution of plant abiotic stress tolerance. *Annals of the New York Academy of Sciences*, 1206, 56–79.

Doroszuk, A., Wojewodzic, M. W., Gort, G. & Kammenga, J. E. (2008). Rapid divergence of genetic variance-covariance matrix within a natural population. *The American Naturalist*, 171, 291–304.

Dugand, R. J., Aguirre, J. D., Hine, E., Blows, M. W. & McGuigan, K. (2021). The contribution of mutation and selection to multivariate quantitative genetic variance in an outbred population of drosophila serrata. *Proceedings of the National Academy of Sciences*, 118, e2026217118.

Eroukhmanoff, F. & Svensson, E. (2011). Evolution and stability of the g-matrix during the colonization of a novel environment. *Journal of evolutionary biology*, 24, 1363–1373.

Falconer, D. S., Mackay, T. F. & Frankham, R. (1996). Introduction to quantitative genetics (4th edn). *Trends in Genetics*, 12, 280.

Gage, J. L., Monier, B., Giri, A. & Buckler, E. S. (2020). Ten years of the maize nested association mapping population: impact, limitations, and future directions. *The Plant Cell*, 32, 2083–2093.

Geiler-Samerotte, K. A., Li, S., Lazaris, C., Taylor, A., Ziv, N., Ramjeawan, C., Paaby, A. B. & Siegal, M. L. (2020). Extent and context dependence of pleiotropy revealed by high-throughput single-cell phenotyping. *PLoS biology*, 18, e3000836.

Geiler-Samerotte, K. A., Zhu, Y. O., Goulet, B. E., Hall, D. W. & Siegal, M. L. (2016). Selection transforms the landscape of genetic variation interacting with hsp90. *PLoS biology*, 14, e2000465.

Hall, M. C., Basten, C. J. & Willis, J. H. (2006). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in mimulus guttatus. *Genetics*, 172, 1829–1844.

Hangartner, S., Lasne, C., Sgrò, C. M., Connallon, T. & Monro, K. (2020). Genetic covariances promote climatic adaptation in australian drosophila. *Evolution*, 74, 326–337.

He, X. & Zhang, J. (2006). Toward a molecular understanding of pleiotropy. *Genetics*, 173, 1885–1891.

Henry, G. A. & Stinchcombe, J. R. (2023). G-matrix stability in clinally diverging populations of an annual weed. *Evolution*, 77, 49–62.

Houle, D., Bolstad, G. H., van der Linde, K. & Hansen, T. F. (2017). Mutation predicts 40 million years of fly wing evolution. *Nature*, 548, 447–450.

Hudson, A. I., Odell, S. G., Dubreuil, P., Tixier, M.-H., Praud, S., Runcie, D. E. & Ross-Ibarra, J. (2022). Analysis of genotype-by-environment interactions in a maize mapping population. *G3*, 12, jkac013.

Jones, A. G., Arnold, S. J. & Bürger, R. (2003). Stability of the g-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution*, 57, 1747–1760.

Jordan, D. M., Verbanck, M. & Do, R. (2019). Hops: a quantitative score reveals pervasive horizontal pleiotropy in human genetic variation is driven by extreme polygenicity of human traits and diseases. *Genome biology*, 20, 1–18.

Juenger, T., Pérez-Pérez, J. M., Bernal, S. & Micol, J. L. (2005). Quantitative trait loci mapping of floral and leaf morphology traits in arabidopsis thaliana: evidence for modular genetic architecture. *Evolution & development*, 7, 259–271.

Kelly, J. K. (2009). Connecting qtls to the g-matrix of evolutionary quantitative genetics. *Evolution: International Journal of Organic Evolution*, 63, 813–825.

Klingenberg, C. P. (2005). Developmental constraints, modules, and evolvability. In: *Variation*. Elsevier, pp. 219–247.

Klingenberg, C. P. (2008). Morphological integration and developmental modularity. *Annual review of ecology, evolution, and systematics*, 39, 115–132.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant, C. & Mott, R. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in arabidopsis thaliana. *PLoS genetics*, 5, e1000551.

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

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*, 94, 203–215.

Lande, R. & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, 1210–1226.

Lea, A., Subramaniam, M., Ko, A., Lehtimäki, T., Raitoharju, E., Kähönen, M., Seppälä, I., Mononen, N., Raitakari, O. T., Ala-Korpela, M. *et al.* (2019). Genetic and environmental perturbations lead to regulatory decoherence. *eLife*, 8, e40538.

Loeschcke, V. (1987). *Genetic constraints on adaptive evolution*. Springer.

Lynch, M., Walsh, B. *et al.* (1998). Genetics and analysis of quantitative traits.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D. & Wolpert, L. (1985). Developmental constraints and evolution: a perspective from the mountain lake conference on development and evolution. *The Quarterly Review of Biology*, 60, 265–287.

McGlothlin, J. W., Kobiela, M. E., Wright, H. V., Mahler, D. L., Kolbe, J. J., Losos, J. B. & Brodie III, E. D. (2018). Adaptive radiation along a deeply conserved genetic line of least resistance in anolis lizards. *Evolution Letters*, 2, 310–322.

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

Meyer, H. V. & Birney, E. (2018). Phenotypesimulator: a comprehensive framework for simulating multi-trait, multi-locus genotype to phenotype relationships. *Bioinformatics*, 34, 2951–2956.

Milocco, L. & Salazar-Ciudad, I. (2022). Evolution of the g matrix under nonlinear genotype-phenotype maps. *The American Naturalist*, 199, 000–000.

Monroe, J. G., Cai, H. & Des Marais, D. L. (2021). Diversity in nonlinear responses to soil moisture shapes evolutionary constraints in brachypodium. *G3*, 11, jkab334.

Muir, C. D., Conesa, M. À., Galmés, J., Pathare, V., Rivera, P., Rodríguez, R. L., Terrazas, T. & Xiong, D. (2022). How important are functional and developmental constraints on phenotypic evolution? an empirical test with the stomatal anatomy of flowering plants.

Opedal, Ø. H., Armbruster, W. S., Hansen, T. F., Holstad, A., Pélabon, C., Andersson, S., Campbell, D. R., Caruso, C. M., Delph, L. F., Eckert, C. G. *et al.* (2023). Evolvability and trait function predict phenotypic divergence of plant populations. *Proceedings of the National Academy of Sciences*, 120, e2203228120.

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

Orr, H. A. (2005). The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, 6, 119–127.

Otto, S. P. (2004). Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271, 705–714.

Paaby, A. B. & Rockman, M. V. (2013). Pleiotropy: what do you mean? reply to zhang and wagner. *Trends in Genetics*, 29, 384.

Pavlicev, M., Cheverud, J. M. & Wagner, G. P. (2009). Measuring morphological integration using eigenvalue variance. *Evolutionary Biology*, 36, 157–170.

Pavlicev, M. & Wagner, G. P. (2012). Coming to grips with evolvability. *Evolution: Education and Outreach*, 5, 231–244.

Pigliucci, M. & Preston, K. (2004). *Phenotypic integration: studying the ecology and evolution of complex phenotypes*. Oxford University Press.

Promislow, D. E. (2004). Protein networks, pleiotropy and the evolution of senescence. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271, 1225–1234.

Rohner, P. T. & Berger, D. (2023). Developmental bias predicts 60 million years of wing shape evolution. *Proceedings of the National Academy of Sciences*, 120, e2211210120.

Saltz, J. B., Hessel, F. C. & Kelly, M. W. (2017). Trait correlations in the genomics era. *Trends in ecology & evolution*, 32, 279–290.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 50, 1766–1774.

Scoville, A., Lee, Y. W., Willis, J. H. & Kelly, J. K. (2009). Contribution of chromosomal polymorphisms to the g-matrix of mimulus guttatus. *New Phytologist*, 183, 803–815.

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, 14, 483–495.

Stearns, F. W. (2010). One hundred years of pleiotropy: a retrospective. *Genetics*, 186, 767–773.

Steiger, J. H. (1980). Testing pattern hypotheses on correlation matrices: Alternative statistics and some empirical results. *Multivariate Behavioral Research*, 15, 335–352.

Steppan, S. J., Phillips, P. C. & Houle, D. (2002). Comparative quantitative genetics: evolution of the g matrix. *Trends in Ecology & Evolution*, 17, 320–327.

Stinchcombe, J. R., Weinig, C., Heath, K. D., Brock, M. T. & Schmitt, J. (2009). Polymorphic genes of major effect: consequences for variation, selection and evolution in arabidopsis thaliana. *Genetics*, 182, 911–922.

Turelli, M. (1988). Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution*, 42, 1342–1347.

Uller, T., Moczek, A. P., Watson, R. A., Brakefield, P. M. & Laland, K. N. (2018). Developmental bias and evolution: A regulatory network perspective. *Genetics*, 209, 949–966.

Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nature genetics*, 50, 693–698.

Waddington, C. H. (1957). *The strategy of the genes*. Routledge.

Wagner, G. P. & Altenberg, L. (1996). Perspective: complex adaptations and the evolution of evolvability. *Evolution*, 50, 967–976.

Wagner, G. P., Kenney-Hunt, J. P., Pavlicev, M., Peck, J. R., Waxman, D. & Cheverud, J. M. (2008). Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature*, 452, 470–472.

Wagner, G. P., Pavlicev, M. & Cheverud, J. M. (2007). The road to modularity. *Nature Reviews Genetics*, 8, 921–931.

Wagner, G. P. & Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics*, 12, 204–213.

Walter, G. M., Aguirre, J. D., Blows, M. W. & Ortiz-Barrientos, D. (2018). Evolution of genetic variance during adaptive radiation. *The American Naturalist*, 191, E108–E128.

Williams, G. C. (2001). Pleiotropy, natural selection, and the evolution of senescence: Evolution 11, 398-411 (1957). *Science of Aging Knowledge Environment*, 2001, cp13–cp13.

Wood, C. W. & Brodie III, E. D. (2015). Environmental effects on the structure of the g-matrix. *Evolution*, 69, 2927–2940.

Zhang, J. & Wagner, G. P. (2013). On the definition and measurement of pleiotropy. *Trends in Genetics*, 29, 383–384.

Appendices

Supplementary Note 1: Distinctions between linkage disequilibrium and tight linkage

Our simulation and analyses account for LD, defined as linkage disequilibrium between markers in a genetic mapping study. Another form of linkage, which is the tight linkage in a genomic context (i.e., perfect linkage, no recombination within pairs of linked loci), is implicitly accounted for when removing major horizontal pleiotropic loci as outliers. We are not able to distinguish between tight linkage and pleiotropy (though these two architectures may differ in maintaining genetic correlation, see Chebib & Guillaume (2021)). These major horizontal pleiotropic loci are either caused by LD within the marker, or by pleiotropy. With respect to distinguishing between LD and horizontal pleiotropy with small effect size loci, polygenic trait correlations have been suggested to be unlikely to arise from chance LD events, as well as several individual horizontal pleiotropic loci (Saltz *et al.*, 2017). Therefore, in essence, our metric dissects the genetic correlation by accouting for allele frequency, linkage equilibrium, as well as horizontal pleiotropy, providing an approach to explore the role of developmental constraint (trait-trait relationship) and vertical pleiotropy (developmental bias in a variant level).

Supplementary Note 2: Additional dataset exhibit qualitatively similar results with yeast morphology dataset

We analyzed an additional data set consisting of vegetative and floral traits measured in a population of *B. rapa L.* in the field and in a greenhouse (Brock *et al.*, 2010). We present results of this analysis based on outlier corrected subset of variants (LD pruned). Fig. S7 shows r_D against r_G. We find that most trait pairs with high r_G also exhibit high r_D, as expected since those trait pairs with high r_G are mostly floral morphological traits, which have long been assumed under developmental integration (Ashman, 1999; Ashman & Majetic, 2006). However, we still found a few exceptions: trait pairs that exhibit low r_D but high r_G. Specifically, for instance, petal length and petal width have a r_G comparable to midpoint length and filament length, petal length and ovary length. However, petal length and petal width have relatively low r_D. This is in contrast to two other pairs of traits, who exhibit higher r_D. Notably, very few length traits appear to have significant r_G with petal width in this dataset (Brock *et al.*, 2010). Juenger *et al.* (2005) also found that, in Arabidopsis, sets of floral organ lengths (petal length, sepal length, long stamen length, pistil length) or organ widths (petal width, sepal width) were highly correlated, while r_G between length and width measures were generally not significant except petal length - petal width. Therefore, we speculate that, instead of r_D as a general cause of r_G, LD or horizontal pleiotropic loci play a critical role in shaping r_G between floral organ lengths and widths. Furthermore, in line with the analyses of yeast morphological data, traits with higher r_D tend to be more conserved in their r_G between conditions: large changes in r_G are more likely to occur for those pairs with

lower r_D(Fig. S8 and Fig. S9).

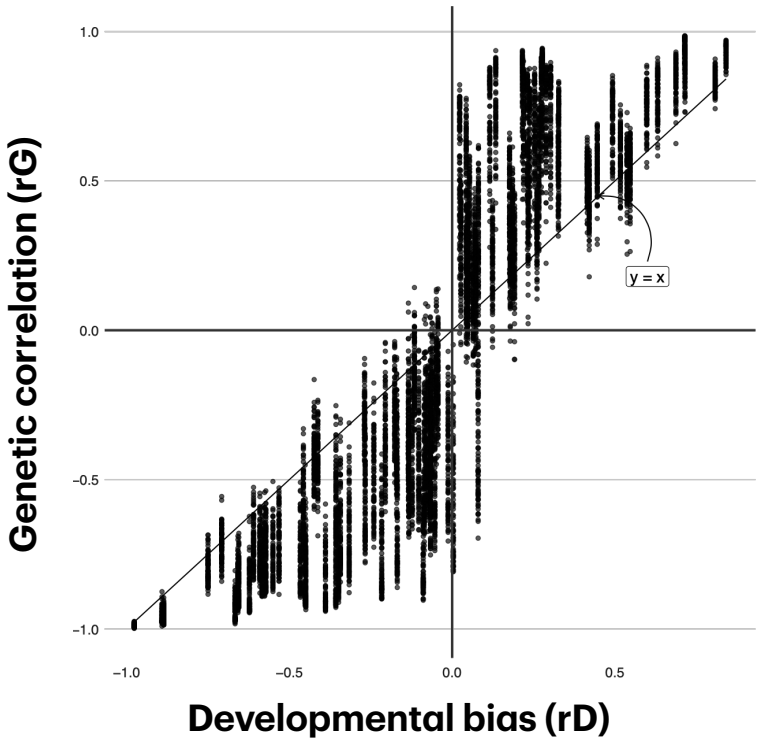

Figure S1: To simulate a pair of traits across populations with different allele frequencies, the additive effects of 1,000 SNPs are simulated using a bivariate normal distribution to generate effect sizes for each locus, with one randomly chosen SNP as a 'major' pleiotropic (additive) locus, and 999 SNPs as 'small' (additive) effects loci which are sampled from a bivariate normal distribution with $\mu = (0, 0)$ and $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. Here, we use a major concordant model, where the effect of major pleiotropic locus in genetic correlation is concordant with the rest of the loci. Next, genotypes of a 100-individual population are generated 100 times with changing allele frequency. Each time, we calculate the genetic and effect size correlation (note that the effect size correlation remains the same). Then, we vary the covariance of bivariate normal distribution and repeat the above step to exhaustively sample different levels of developmental bias.

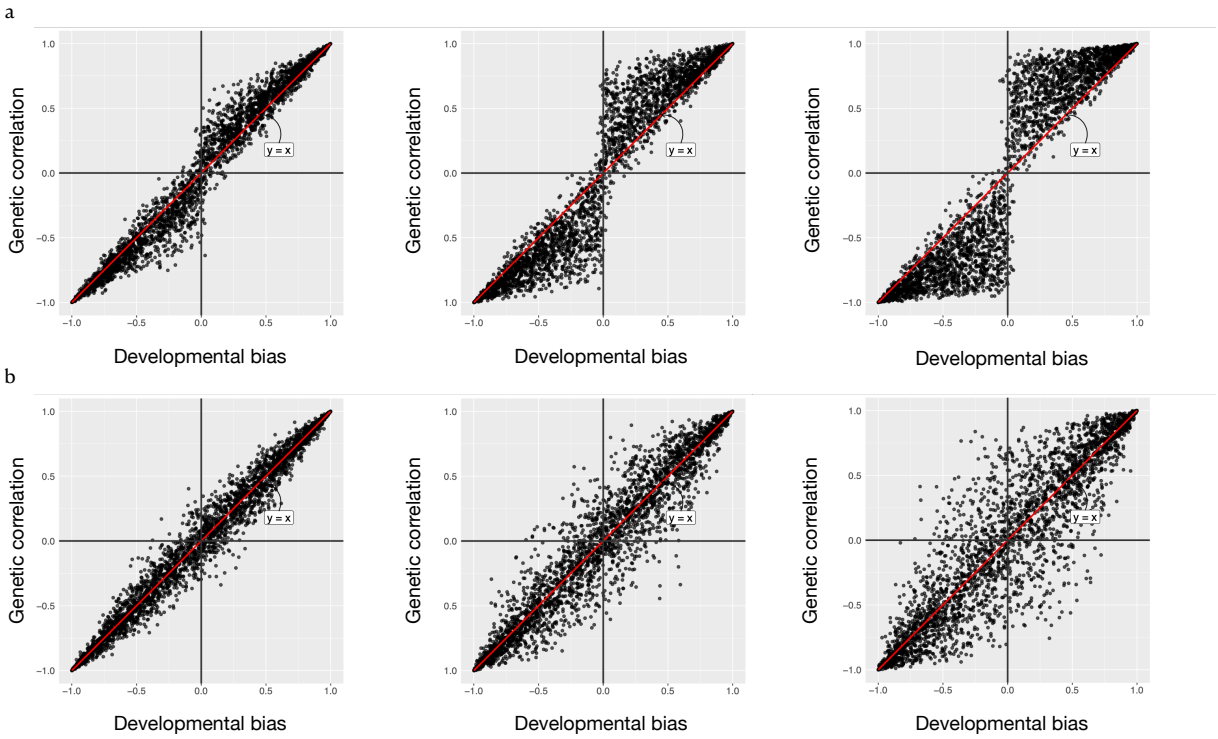

Figure S2: Reproduces the scatter plot between developmental bias (r_D) and genetic correlation (r_G) in Fig. 2, but here using a bivariate normal distribution to generate effect sizes for each loci. Each point represents a pair of traits. Given a population, we simulated 3,000 pairs of traits using a model under which each trait pair consists of one randomly chosen SNP with a 'large' pleiotropic (additive) effect, and 999 SNPs with 'small' (additive) effects which are sampled from a bivariate normal distribution with $\mu = (0, 0)$ and $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. ρ is sampled from a uniform distribution $(-1, 1)$. **a.** simulations under a concordant model, where effect of the large pleiotropic loci on genetic correlation is concordant with the rest of genetic background. **b.** simulations without concordant assumption, where large pleiotropic loci has either concordant or antagonistic effect with the genetic background. From left to right, effect size of the large pleiotropic loci increases.

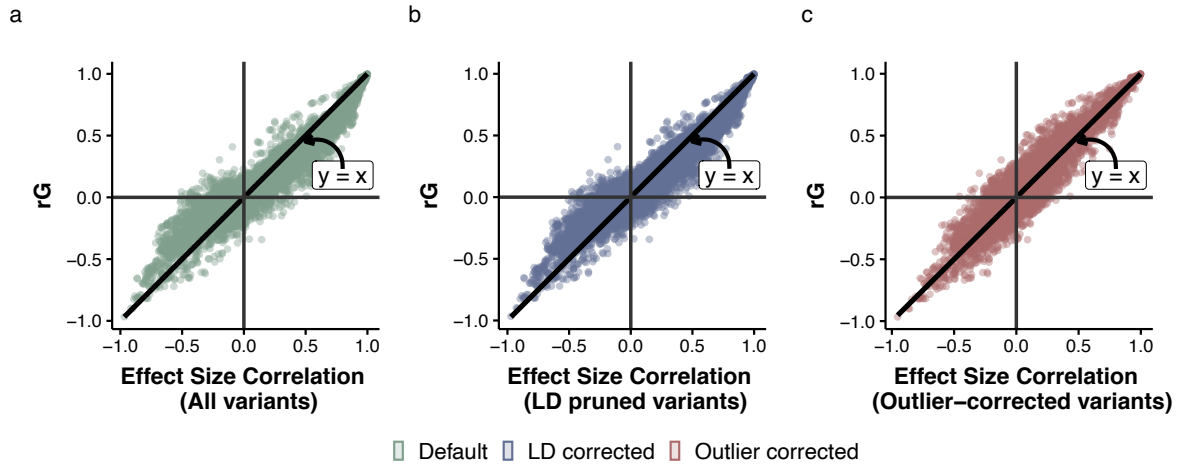

Figure S3: The relationship between r_D and r_G from our re-analysis of phenotypes measured in 374 recombinant strains of yeast cells (Geiler-Samerotte *et al.*, 2020). Each dot in the scatter plot represents, for a given pair of traits in the yeast dataset which consists of 167 traits, the correlation of additive effect (on the x axis) and the genetic correlation (r_G, on the y axis). We consider three settings, summarized in Fig. 4. **a.** All loci across the genome. **b.** LD pruned variants. **c.** Outlier corrected variants. In **a.** and **b.**, effect size correlations are stronger than genetic correlations (Wilcoxon signed-rank test, p-value $<$ 2.2e-16). Conversely, in **c.**, effect size correlations are not stronger than genetic correlations (Wilcoxon signed-rank test, p-value = 0.8875) and the bulk of data are more evenly distributed around the unity line)

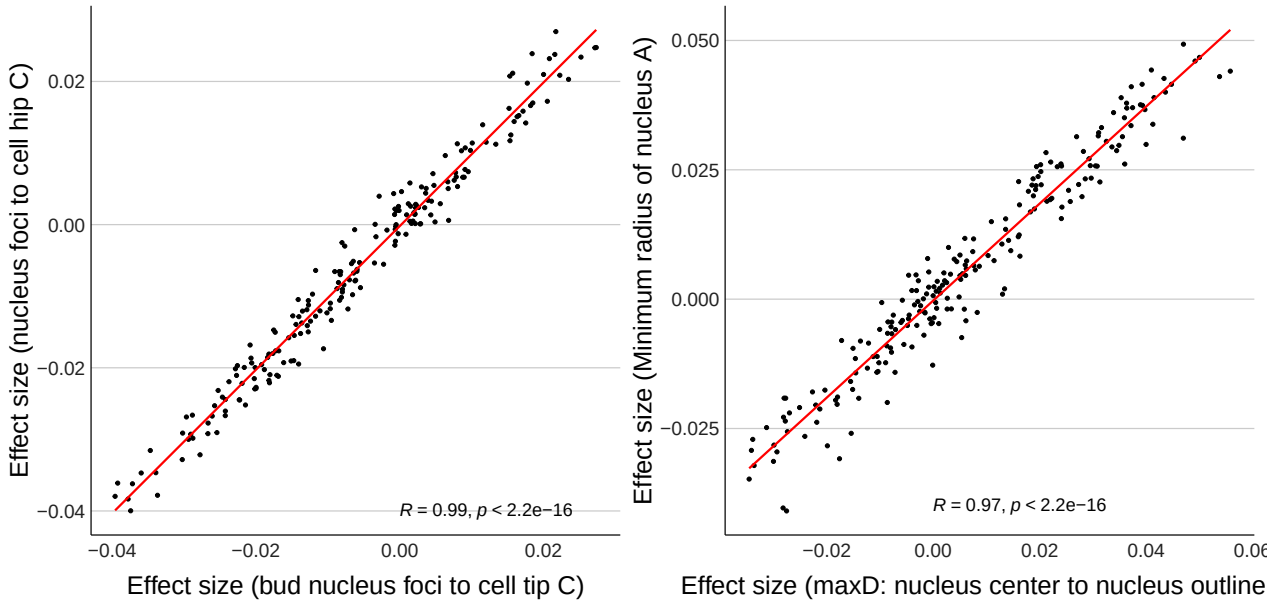

Figure S4: Two example trait pairs in the yeast morphology dataset (Geiler-Samerotte *et al.*, 2020) showing exceptionally strong r_D between traits. Each point represents additive effect for a single locus. These trait pairs demonstrate strong inherent redundancy.

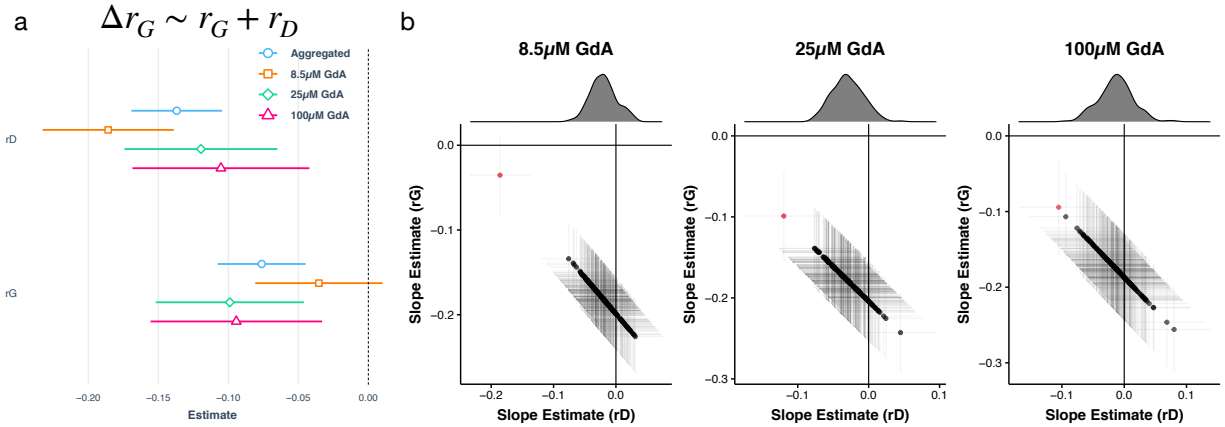

Figure S5: Statistical test showing developmental bias (r_D) can predict the changes of genetic correlation (r_G). **a** Linear regression results for $\Delta r_G \sim r_D + r_G$. Estimates of coefficient and their 95% confidence intervals under four conditions for r_D and r_G are shown. The environmental conditions are three concentrations of geldanamycin (GdA) plus an aggregated condition with all three concentration data. GdA is a small-molecule inhibitor that binds the ATP-binding site of the chaperone Hsp90, thus rendering it unable to perform its cellular function. **b** Since r_G and r_D are highly correlated, which might cause multi-colinearity problems during regressions, we perform additional analyses by simulating a null model (black dots): $\Delta r_G \sim r_n + r_G$, where r_n is sampled from the bivariate normal distribution with covariance 0.939, conditioning on r_G. The coefficient estimates of r_D (red dots) in **b** under all conditions are at the tail of distribution in null expectation towards stronger slope estimates, suggesting r_D provides additional information in predicting Δr_G.

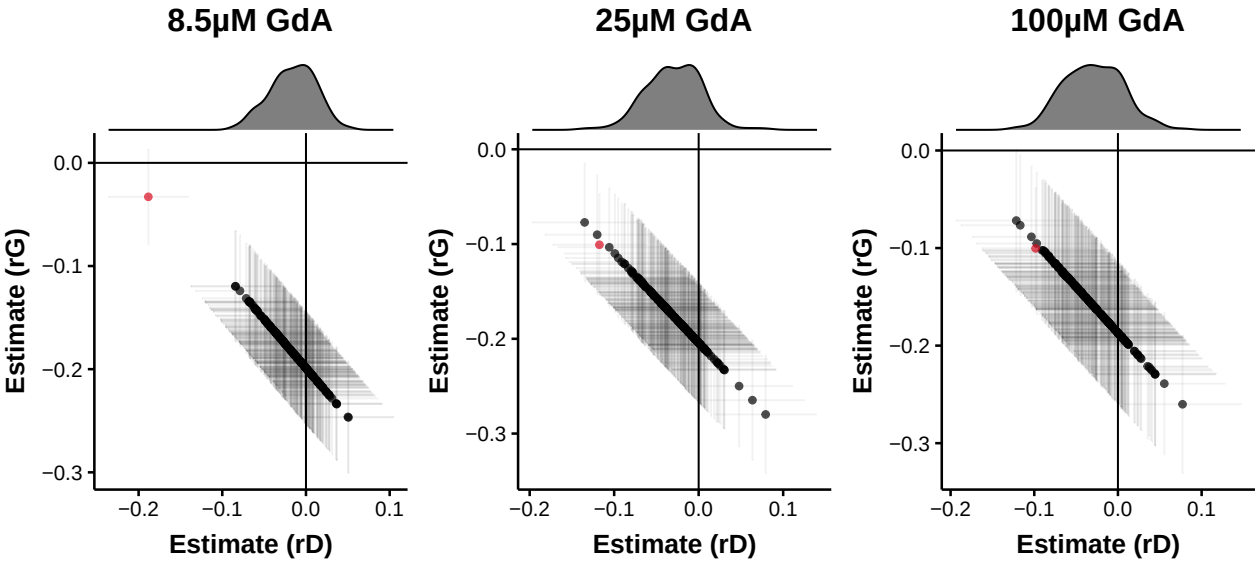

Figure S6: Statistical test showing developmental bias (r_D) can predict the changes of genetic correlation (r_G), similarly with Fig. S5b. Estimates of coefficient and their 95% confidence intervals under four conditions for r_D and r_G in a linear regression $\Delta r_G \sim r_D + r_G$ are shown in red dots. Since r_G and r_D are highly correlated, which might cause multi-colinearity problems during regressions, we perform additional analyses by simulating a null model (black dots): $\Delta r_G \sim r_n + r_G$, where rn is generated by r_G with Gaussian noise. Qualitatively similar results with (Fig. S5b) are observed

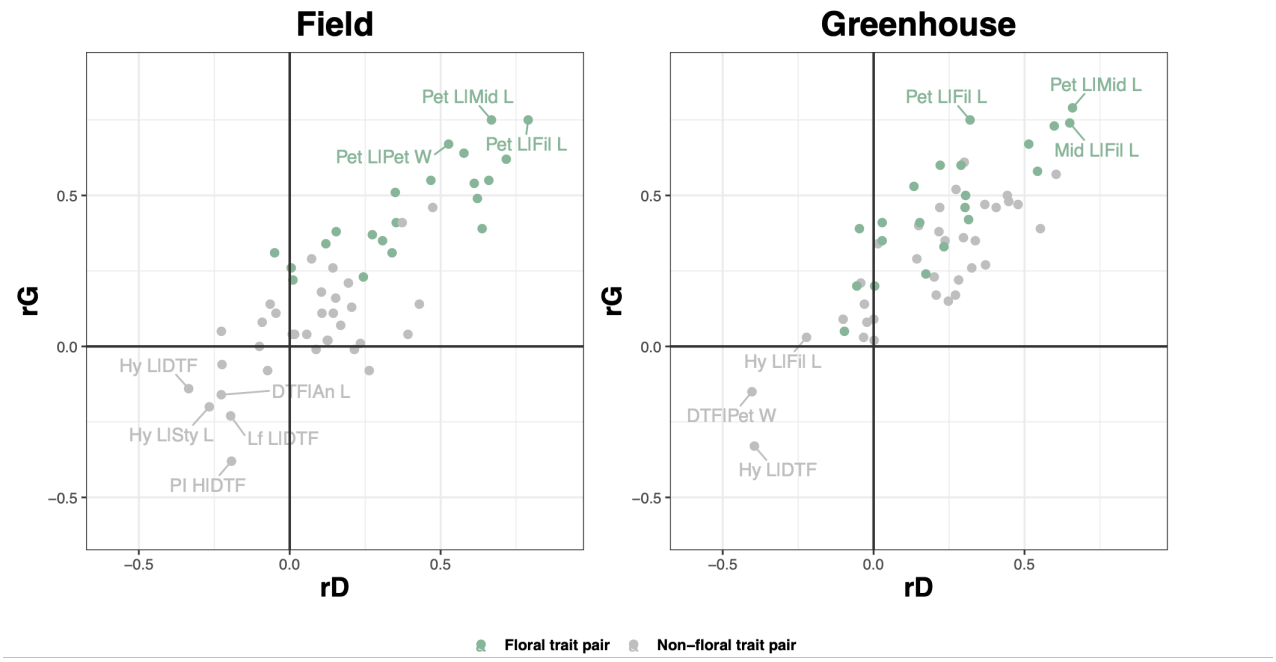

Figure S7: A *Brassica* dataset with 11 floral, vegetative, and phenology traits (Brock *et al.*, 2010) was analysed similarly to Fig. 3, under two environmental conditions (field and greenhouse). The loci used to calculate the effect size correlation are LD pruned and outlier corrected. Green dots denote pairs of traits that are both floral.

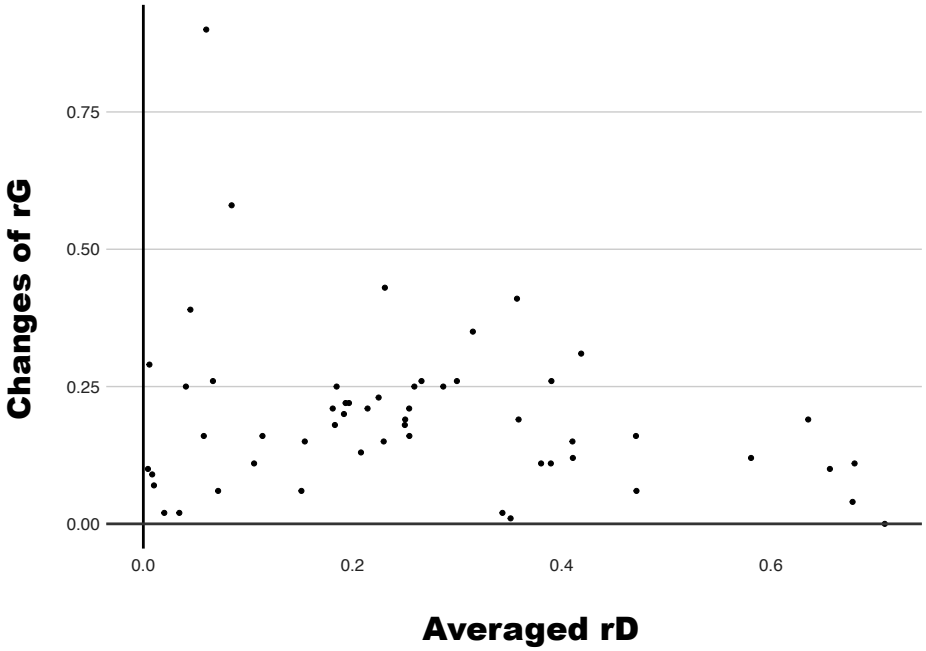

Figure S8: A *Brassica* dataset with 11 floral, vegetative, and phenology traits (Brock *et al.*, 2010) under two environments (field and greenhouse) was analysed. Trait pairs with stronger r_D exhibit more stable genetic correlations across two environmental conditions. The x axis is the averaged developmental bias across two environments. y axis shows the absolute difference of genetic correlation for a given pair.

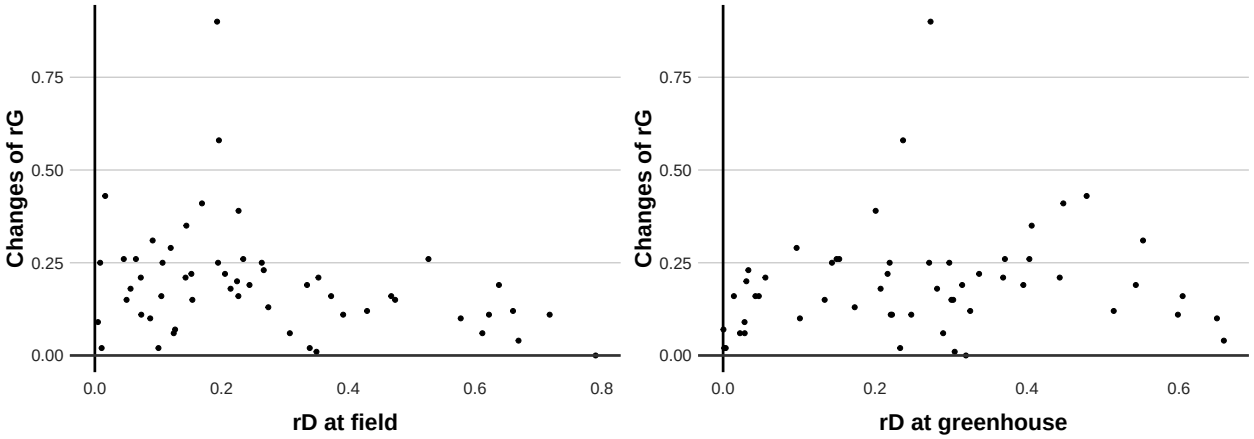

Figure S9: A *Brassica* dataset with 11 floral, vegetative, and phenology traits (Brock *et al.*, 2010) under two environments (field and greenhouse) was analysed. Trait pairs with stronger r_D exhibit more stable genetic correlations across two environmental conditions.The x axis is the developmental bias in two environments, respectively. y axis shows the absolute difference of genetic correlation for a given pair.