Differential Regulation of Cryptic Genetic Variation Shapes the Genetic Interactome Underlying Complex Traits

Anupama Yadav[1], Kaustubh Dhole[1], and Himanshu Sinha[1,2,3,*]

[1]Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai, India

[2]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai, India

[3]Initiative for Biological Systems Engineering, Indian Institute of Technology Madras, Chennai, India

*Corresponding author: E-mail: sinha@iitm.ac.in.

Accepted: October 21, 2016

Abstract

Cryptic genetic variation (CGV) refers to genetic variants whose effects are buffered in most conditions but manifest phenotypically upon specific genetic and environmental perturbations. Despite having a central role in adaptation, contribution of CGV to regulation of quantitative traits is unclear. Instead, a relatively simplistic architecture of additive genetic loci is known to regulate phenotypic variation in most traits. In this paper, we investigate the regulation of CGV and its implication on the genetic architecture of quantitative traits at a genome-wide level. We use a previously published dataset of biparental recombinant population of *Saccharomyces cerevisiae* phenotyped in 34 diverse environments to perform single locus, two-locus, and covariance mapping. We identify loci that have independent additive effects as well as those which regulate the phenotypic manifestation of other genetic variants (variance QTL). We find that whereas additive genetic variance is predominant, a higher order genetic interaction network regulates variation in certain environments. Despite containing pleiotropic loci, with effects across environments, these genetic networks are highly environment specific. CGV is buffered under most allelic combinations of these networks and perturbed only in rare combinations resulting in high phenotypic variance. The presence of such environment specific genetic networks is the underlying cause of abundant gene–environment interactions. We demonstrate that overlaying identified molecular networks on such genetic networks can identify potential candidate genes and underlying mechanisms regulating phenotypic variation. Such an integrated approach applied to human disease datasets has the potential to improve the ability to predict disease predisposition and identify specific therapeutic targets.

Key words: cryptic genetic variation, phenotypic buffering, variance QTL, gene–environment interaction, genetic networks.

Introduction

Understanding how variation at genetic level bears on phenotypic variation is the fundamental question in biology. Multiple mapping studies in laboratory model systems, domesticated populations and human populations have identified a predominant additive regulation of phenotypic variation, i.e., despite presence of high genetic diversity, only a few independently acting loci result in the observed phenotypic variation (Mackay 2001; Mackay et al. 2009). A reason behind such an additive architecture is the focus on the identification of loci with main effects, i.e., effects averaged over phenotype of all other loci (Mackay 2014). Attempts to identify epistasis have been limited by the lack of power in most organisms (Phillips 2008). The studies that have analyzed epistasis show minimal effects of genetic interactions on phenotypic variation (Bloom et al. 2015). Paradoxically, despite limited discovery of epistasis, the identified additive loci show high genetic background dependence (Mackay 2001; Mackay et al. 2009), questioning their independent effects. In parallel, a highly interconnected non-linear architecture of genetic networks is emerging from the numerous genome-wide molecular studies (Vidal et al. 2011). These studies in various organisms including yeast, fly, worms, mice, and humans have identified a complex interactome constituting the genotype–phenotype map. A salient feature of these interactomes is the presence of regulatory hubs, which cross talk with other genes to mediate phenotype. This argues for existence of such a

multi-layered architecture of the genetic variants that regulate phenotypic variation. In this study, we ask why only a subset of the diverse genetic variation has been associated with phenotypic variation. Is such an additive architecture a true reflection of the genotype–phenotype map or a result of the bias in the mapping approaches?

To understand the architecture of genetic variation, it is important to take into account various evolutionary processes that result in accumulation of this genetic variation. An adapted population is not genetically constant, instead it continues to accumulate genetic variants phenotypically neutral in that selection condition (Gibson and Dworkin 2004). In a scenario of a change of genetic background or environment, these accumulated variants can manifest phenotypically, i.e., show their effects, thereby facilitating adaptation. This genetic potential is termed cryptic genetic variation (CGV). Various studies have investigated regulation of CGV at a single locus level. Waddington's initial observations in *Drosophila* later lead to identification of genes with the ability to buffer the phenotypic effects of other genetic variants or CGV in various systems—*Ultrabithorax* (Gibson et al. 1999), *EGFR* (Dworkin et al. 2003) in *Drosophila*, *IRA2* (Taylor and Ehrenreich 2015) in yeast and, the most commonly studied gene, *HSP90* (Rutherford and Lindquist 1998) in various model systems. Upon genetic or environmental perturbation, the buffering ability of these genes is altered, resulting in phenotypic manifestation of CGV. However, these few examples of regulators of CGV cannot account for its proposed central role in adaptation and speciation (Gibson and Dworkin 2004). At a population level, CGV refers to genetic variation that is masked through epistatic effects and released in the presence of specific polymorphisms (Hermisson and Wagner 2004; Siegal and Leu 2014). However, the mainly additive nature of causal genetic variation in mapping studies challenges the role of loci regulating CGV in phenotypic variation. At a molecular level, high genetic crosstalk resulting in redundant pathways (Wu and Lai 2015) has been proposed to be a compensatory mechanism to buffer CGV (Costanzo et al. 2010) indicating its pervasive role in shaping the genetic architecture. Therefore, in order to understand the impact of CGV on phenotypic variation, it is important to study its regulation at a genome-wide level.

In this paper, we investigate genetic architecture regulating phenotypic variation by employing QTL mapping to study regulation of CGV. We ask the following questions: can we identify loci whose alleles have differential effects on the phenotypic manifestation of other genetic variants, i.e., CGV? If such loci exist, how do they compare to the conventionally identified loci and how are they distributed across diverse environments? Finally, like the highly interconnected molecular interactome, do the genetic variants also show high interconnectivity and how does this impinge on regulation of CGV?

Linkage mapping in segregants of a large biparental population will allow comparison of the effect of the two divergent alleles of a locus on thousands of genetic variants, i.e., CGV. Whereas conventional quantitative trait locus (QTL) mapping identifies alleles that show significant difference in population mean, their effects on population variance are not considered (Rönnegård and Valdar 2012). A significant difference in mean of the two alleles indicates difference in mean phenotype averaged over other genetic variants. Whereas, a difference in variance would indicate that while rest of the genetic variants are phenotypically neutral or buffered in the presence of one allele resulting in lower variance, the other allele allows their phenotypic manifestation resulting in higher variance (Lempe et al. 2013). The term to describe these variance-regulating loci, variance QTL (vQTL) was first introduced by Rönnegård and Valdar (2011). Whereas vQTL with significant effect on variation have been identified, viz. *MOT1* (Forsberg et al. 2015) and *nFT* (Lee et al. 2014) in plants, these loci tend to be small effect when compared with conventional QTL (Shen et al. 2012). In addition, while vQTL mapping has been proposed to be a predictor of prevalence of gene-gene interactions (Paré et al. 2010; Rönnegård and Valdar 2012), thereby supporting our hypothesis of differential regulation of CGV by these loci, empirical evidence for this remains scarce.

We used linkage mapping to identify loci that regulate CGV in a biparental yeast population grown in diverse environments. These genetically diverse parental strains (BY and RM11) have accumulated large number of genetic variants over the course of their evolutionary trajectories (Liti et al. 2009; Bloom et al. 2013). We perform single locus mean and variance QTL mapping and compare their distribution in diverse environments. We then perform across environment covariance analysis of the identified loci to study the environment specific pleiotropy of the regulation of CGV and its impact on gene–environment interactions (GEI). Please note that from this point onwards in this paper, pleiotropy refers to environment specific pleiotropy, i.e., effect of a locus in multiple environments. In the end, we perform a two-locus analysis to detect genetic loci whose interactions affect mean, variance or both and identify genetic networks regulating CGV. Our results demonstrate that vQTL mapping can be used to identify the regulators of CGV. We show that environment is the primary determinant of regulation of CGV, such that it is buffered in most environments and is released only in certain environments. While most large effect pleiotropic loci affect CGV in these select environments, we identify environment specific genetic and molecular networks, which work in a redundant manner to buffer the CGV. We believe that such networks are a part of genetic architecture of all quantitative traits and these networks are revealed only upon certain genetic and environmental perturbations.

Methods

In this paper, for clarity, we have used following terms to define various mapping techniques used: "mQTL", mean QTL, is used to describe conventional QTL mapping based on F-statistic comparison; "vQTL" for variance QTL mapping based on BF-statistic comparison; "mvQTL" for a locus which has a significant F- and BF-statistic; "2-mQTL" denotes pair of loci with significant two-way F-statistic comparison; "2-vQTL" with two-way BF-statistic comparison; and "2-mvQTL" with two-way F- and BF-statistic comparison.

Dataset

The raw growth data analyzed in this study was derived from a study by Bloom et al. (2013), in which the experimental procedures are described in detail. The data we used was generated for 1,008 segregants derived from a cross between *S. cerevisiae* strains BY (a laboratory strain) and RM11-1a (a wine isolate, indicated as RM). These segregants were grown in 46 different environments and phenotyped for colony size of which 34 environments were considered in this study (supplementary table S1, Supplementary Material online). A total of 11,623 markers were considered.

mQTL and vQTL Mapping

The mQTL mapping was carried out as described previously (Bhatia et al. 2014). In brief, the R/qtl package (Broman et al. 2003; Broman and Sen 2009) was used to identify QTL separately for colony size in each environment. mQTL were identified using the LOD score, which is the $\log_{10}$ of the ratio of the likelihood of the experimental hypothesis to the likelihood of the null hypothesis (Broman and Sen 2009). An interval mapping method using R/qtl "scanone" function was used to compute this LOD score using the Haley–Knott regression algorithm (Broman et al. 2003).

The following formula was used to calculate the F-score, which was further used to derive the LOD score. At a particular marker, let segregant i's phenotypic value be y_{ij} where j can take two values ($j = 1$: BY allele and $j = 2$: RM allele).

$$F = \frac{\sum_{j=1}^{k} n_j \left(\overline{y}_j - \overline{y}\right)^2 / (k-1)}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(y_{ij} - \overline{y}_j\right)^2 / (n-k)}$$

Here, N is the total number of segregants, n_1 and n_2 are the number of segregants having the BY and RM allele, respectively ($k = 2$) and y_i is the genotypic mean of allele j.

Let df denote the degrees of freedom ($df = 1$ for a backcross and $df = 2$ for an intercross). The LOD score is accordingly derived as follows:

$$LOD = \frac{n}{2} \log_{10} \left[F \left(\frac{df}{n - df - 1} \right) + 1 \right]$$

Under the null hypothesis, there is no significant difference in the means at the marker under consideration whereas under the alternative hypothesis, there is a presence of an mQTL.

To estimate the difference in phenotypic variance between the two genotypic groups, i.e., to identify vQTL in each environment, the standard Brown–Forsythe (BF) statistic (Rönnegård and Valdar 2011; Lee et al. 2014) and the corresponding LOD score were calculated for each genetic marker in each environment (supplementary file S1, Supplementary Material online). The BF-test is equivalent to an F-test performed on the deviations of the phenotypic values from their respective genotypic medians (or the means). Hence, under the alternative hypothesis, the phenotypes of the two alleles reveal a difference in the variance.

At a particular marker, let z_{ij} be the absolute deviation of segregant i's phenotypic value y_{ij} from its genotypic mean $\tilde{y}_j$ where, j can take two values ($j = 1$: BY allele and $j = 2$: RM allele).

$$z_{ij} = |y_{ij} - \tilde{y}_j|$$

Then BF-statistic for that marker can be computed as follows:

$$F = \frac{(N-p)}{(p-1)} \frac{\sum_{j=1}^{p} n_j \left(\tilde{z}_j - \tilde{z}_{..}\right)^2}{\sum_{j=1}^{p} \sum_{i=1}^{n_j} \left(z_{ij} - \tilde{z}_j\right)^2}$$

Here, N is the total number of segregants, n_1 and n_2 are the number of segregants having the BY and RM allele, respectively ($p = 2$). In order to estimate the effects of vQTL in the same order as in QTL, LOD scores were computed as described previously (Broman and Sen 2009).

To establish the statistical significance of the putative mQTL and vQTL, P values were computed using a genome-wide permutation test of 1,000 permutations, where the null distribution consisted of the highest genome-wide LOD score obtained from each permutation (Broman et al. 2003). A marker with LOD score > 3.0 and a permutation $P < 0.1$ was considered significant. In order to account for multiple comparisons performed on this dataset, we used permutation testing which represents a more robust substitute to various multiple testing correction algorithms (Camargo et al. 2008). In brief, permutations disrupt the genotype–phenotype association and then test the chances of identification of a particular locus as significant. This is independent of any other test performed on the dataset. In order to test the association between permutation values for F- and BF-tests, we calculated

correlation of permutation P values of the two tests for all single QTL and identified a non-significant correlation of $r^2 = -0.143$.

Comparing mQTL and vQTL and Estimating Pleiotropy

To compare mQTL and vQTL and estimate pleiotropy, we divided the genome into 50 kb non-overlapping regions. A 50 kb region containing significant mQTL and vQTL in the same environment was considered an mvQTL. Similarly a 50 kb region containing two or more mQTL or vQTL significant in different environments was considered a pleiotropic locus. Within one type of mapping in an environment, we did not identify two peaks within a 50 kb locus. This is also tested independently in the original study, Bloom et al. (2013), which did not identify two peaks closer than 50 kb in the same environment. Furthermore, we observed that for significant loci any linkage degenerated beyond a 50 kb interval independent of the LOD score. Therefore, a 50 kb region was sufficient to compare loci pleiotropic across environments. We compared 40 and 50 kb intervals and found no difference in our results. We have therefore used a universal QTL interval of 50 kb. We would like to note that linked loci can have an effect in the same environment as well, as reported previously (Steinmetz et al. 2002) and distinguishing the effect of such loci across environments is beyond the scope of linkage mapping and this paper.

Calculating Correlation between Mean and Variance

To calculate environment specific correlation, difference in the mean (mean_BY − mean_RM) and difference in the variance (var_BY − var_RM) was calculated for all significant loci within each environment. Pearson's correlation between these two parameters was then calculated for each environment.

Covariance across Environmental Pairs

To assess the differential covariance of a locus across multiple pairs of environments, we considered the collated set of pleiotropic loci for our study (supplementary table S3, Supplementary Material online). To quantify the differential covariance across a pair of environments, a Deming regression was calculated between the phenotype values of the chosen pair of environments for each allele, using R package "mcr". Deming regression, which minimizes errors in multiple dimensions simultaneously, served as a suitable measurement error model for assessing buffering across two or more environments. Unlike simple least squares regression, Deming regression accounts for deviations in observations on both the x- and the y-axis. Assuming (y_i, x_i) are measured observations of the "true" values (y_i^*, x_i^*), which lie on the regression line.

$$y_i = y_i^* + \varepsilon_i$$

$$x_i = x_i^* + \eta_i$$

where, errors ε and η are independent and the ratio of their variances is the estimate $\delta = \frac{\sigma_\varepsilon^2}{\sigma_\eta^2}$. Deming regression seeks to find the line of best fit: $y^* = \beta_0 + \beta_1 x^*$, such that the weighted sum of squared residuals is minimized.

$$\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\left(\left(y_i - \beta_0 - \beta_1 x_i^*\right)^2 + \delta\left(x_i - x_i^*\right)^2\right)$$

Because both the phenotypic values (y_i, x_i) are normally distributed across their respective axes, Deming regression, serves as a suitable measurement-error model for assessing buffering across two environments. Two Deming regression models were fitted for each environment pair corresponding to the two alleles.

A Student's t-test was performed between the deviations of the phenotypic values from the Deming fit of the BY and RM alleles ($P < 0.05$). The sign of the t-test value indicated which of the alleles had higher mean of deviations, i.e., lower covariance. A positive t-test value denotes that the mean deviation of BY allele was greater than that of the RM allele and vice versa. All the significant loci and their corresponding environmental pairs along with mean deviations of the alleles and independent mean and variance of both the alleles and both the environments are given in supplementary table S3, Supplementary Material online. A locus with significantly different allelic covariance across more than 15 environmental pairs was considered to have pleiotropic effect on covariance. For each of these loci, total number of positive and negative t-test statistic were compiled. A Fisher's Exact test of random distribution of positive and negative t-test statistic was done for each locus and corrected for the total number of environmental pairs considered. A $P < 0.05$ was considered significant.

2-mQTL and 2-vQTL Mapping

A 2-mQTL interaction occurs when the effect of an mQTL at a single locus depends on another locus. We used a two-locus mapping technique described previously to test the interactions (Bhatia et al. 2014). LOD scores were computed for pairwise comparisons among a set of selected loci, instead of doing a whole genome analysis. To increase power to identify 2-mQTL and 2-vQTL interactions, for each environment, genetic loci significant in either QTL or vQTL or mvQTL mapping were collated for each environment (supplementary table S1, Supplementary Material online). For the 2-mQTL mapping, P values were computed using a permutation test (10,000 permutations), where the null distribution consisted of the highest LOD score obtained among all pairwise comparison for each permutation of the phenotype. Interactions with permutation $P < 0.1$ were considered significant. The following hypotheses were compared:

$$H_I : y_i = \mu + \beta_1 g_{1i} + \beta_2 g_{2i} + \gamma g_{1i} g_{2i} + \epsilon_i$$

$$H_A : y_i = \mu + \beta_1 g_{1i} + \beta_2 g_{2i} + \epsilon_i.$$

Here, g_{1i} and g_{2i} are the binary variables that specify the genotype at the two loci, μ, β_1, β_2, and γ are inferred from the data using maximum likelihood. The parameters β_1 and β_2 quantify the individual effect of each QTL, and γ quantifies the effect of the 2-mQTL interaction. See supplementary file S1, Supplementary Material online, for details of calculation of the interaction LOD (LOD_i) score.

Apart from the 2-mQTL mapping described above, we mapped variance-controlling interactions by mapping 2-vQTL interactions, which occurs when the phenotypic variance at one locus depends on the genotype at another locus. The hypothesis testing for 2-vQTL interactions was done in the same way as 2-mQTL interactions except that instead of scaled values of colony size as the phenotypic values, deviations from the mean were used as the phenotype for each allelic combination (Method for vQTL mapping). In order to compare 2-mQTL and 2-vQTL across environments, loci within a 50 kb linkage interval were considered the same.

Identifying Molecular Networks

For each genetic network independently, we compiled the genes present in the constituent loci using Bloom et al. (2013) data. Interactions between these gene lists were identified using STRING (http://string-db.org; last accessed May 27, 2016) (Szklarczyk et al. 2015) with following settings: interaction sources selected were experimental evidence, curated databases and co-expression with interaction score of medium confidence (0.4).

Results

A Non-Uniform Distribution of Loci Affecting Mean and Variance

Using a previously published dataset (Bloom et al. 2013) of a recombinant haploid population generated from a biparental cross between a laboratory strain BY and a vineyard isolate RM, we carried out single locus linkage mapping to identify genetic loci that showed an effect on the mean (mQTL) and variance (vQTL) of colony size variation. This mapping was done independently in 34 diverse environments, ranging from different carbon sources to oxidative and DNA damaging stresses. mQTL were estimated using F-test, whereas vQTL were estimated using Brown–Forsythe (BF) test (fig. 1*A*, Methods). Significance was determined by performing 1,000 permutations per marker and a locus with LOD score > 3 and permutation $P < 0.1$ was termed significant. If the peaks of mQTL and vQTL in one environment were within 50 kb interval, the locus was termed mvQTL, i.e., this locus that had an effect on both mean and variance. The environments considered had a high broad sense heritability ($H^2 > 0.5$) estimated using replicates of more than 500 segregants in most cases (Bloom et al. 2013). This robustness of the replicates ensured that the phenotype of segregants was robust and the differential variance was not a result of technical noise in phenotype estimation.

A total of 296 significant loci (LOD score > 3.0; permutation $P < 0.1$, supplementary table S1, Supplementary Material online) were identified across the 34 environments, of which 72% (212) behaved solely as mQTL, 6% (19) uniquely as vQTL and the remaining 22% (65) as mvQTL (fig 1*B*). Two main conclusions emerge from these results. First, almost three-fourths of the loci are mQTL, i.e., they have independent additive effects and only one-fourth of the loci result in a difference in variance. Furthermore, among the loci resulting in difference in variance, a substantially higher number of mvQTL were identified compared with vQTL. This indicated that in most cases variance heterogeneity is accompanied by a difference in mean. Even though we did not identify many additional loci compared with the original study (Bloom et al. 2013), using two methods of QTL mapping (mQTL and vQTL), we classified the identified loci into different categories based on their allelic variance heterogeneity. Second, these vQTL and mvQTL loci are not uniformly distributed in all environments, instead, they are more prevalent in a few select environments (fig. 1*B* and *D*). If the difference in variance were a statistical effect, then vQTL and mvQTL would have been present with similar frequency across all environments without the observed distribution bias. Therefore, this observation overrules a purely statistical origin of mvQTL without any biological implication. Additionally, we observe a strong negative correlation ($r^2 = -0.9$) between frequencies of mQTL and mvQTL, indicating exclusivity of presence of either mQTL or mvQTL in an environment, but not both. In other words, two types of genetic architectures regulating phenotypic variation appear to exist in different environments, one consisting predominantly of mQTL and other of mvQTL.

Environment Determines Association between Population Mean and Variance

A higher frequency of mvQTL compared with vQTL indicates that in most cases increase in variance is accompanied by a shift in the mean resulting in a higher or a lower average phenotype. We next asked what determined the higher or lower direction of the shift in mean; was it is a property of the locus or determined by the environment. A positive direction of release of variance would mean that an increase in variance is associated with an increase in mean phenotype, whereas a negative direction of release of variance would mean that an increase in variance is associated with a decrease in phenotypic mean (fig. 2*A*). In order to test if this directionality of release of variance was a property of the locus, as a first step, pleiotropic loci that affected both mean and variance

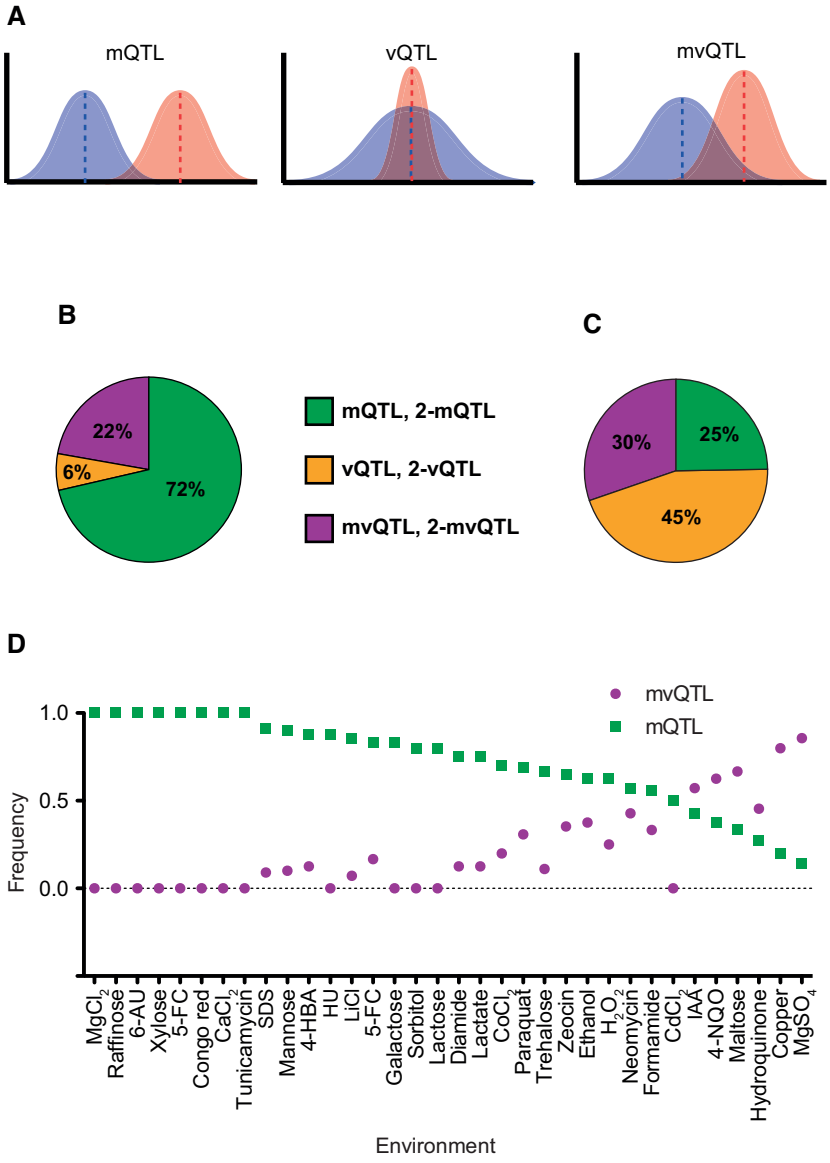

Fig. 1.—Distribution of mQTL, vQTL and vQTL. (*A*) Schematic showing the three categories of QTL mapped. mQTL has a significantly different allelic mean but a non-significant difference in allelic variance; vQTL has a non-significant difference in mean but significantly different variance; mvQTL has both significantly different allelic mean and variance. (*B*) Distribution of the three categories of QTL mapped, mQTL (green), vQTL (orange), and mvQTL (dark pink). (*C*) Distribution of the three categories of two-locus interaction mapped, 2-mQTL (green), 2-vQTL (orange), and 2-mvQTL (dark pink). (*D*) Distribution of loci mapped in each environment as mQTL (green) and mvQTL (dark pink) in the segregating population. The *y*-axis is proportion of mQTL or mvQTL in each category. The *x*-axis is different environments arranged by decreasing proportion of mQTL. Permutation *P* value cut off <0.1. See supplementary table S1, Supplementary Material online, for details.

across multiple environments were identified. As described previously, fewer loci behaved as mvQTL (65/296) compared with mQTL (212/296), with most of the mvQTL being environment specific (supplementary table S1, Supplementary Material online). Pearson correlation was calculated between the mean and the variance of the alleles of each locus across different environments. Little correlation was observed between the allelic mean and variance of these pleiotropic loci across environments. For example, a chrXV (140,012) locus had a pleiotropic effect on mean and variance across eight environments. Whereas the BY allele had a higher mean in $MgCl_2$ and LiCl, and RM allele in $MgSO_4$, the RM allele had a higher variance in all these three environments showing a poor correlation ($r^2=0.3$) between the mean and the variance (fig. 2*B*, supplementary table S2, Supplementary Material online). A similar lack of association between the mean and the variance was observed for other loci as well. On the other hand, a strong association was observed between the mean

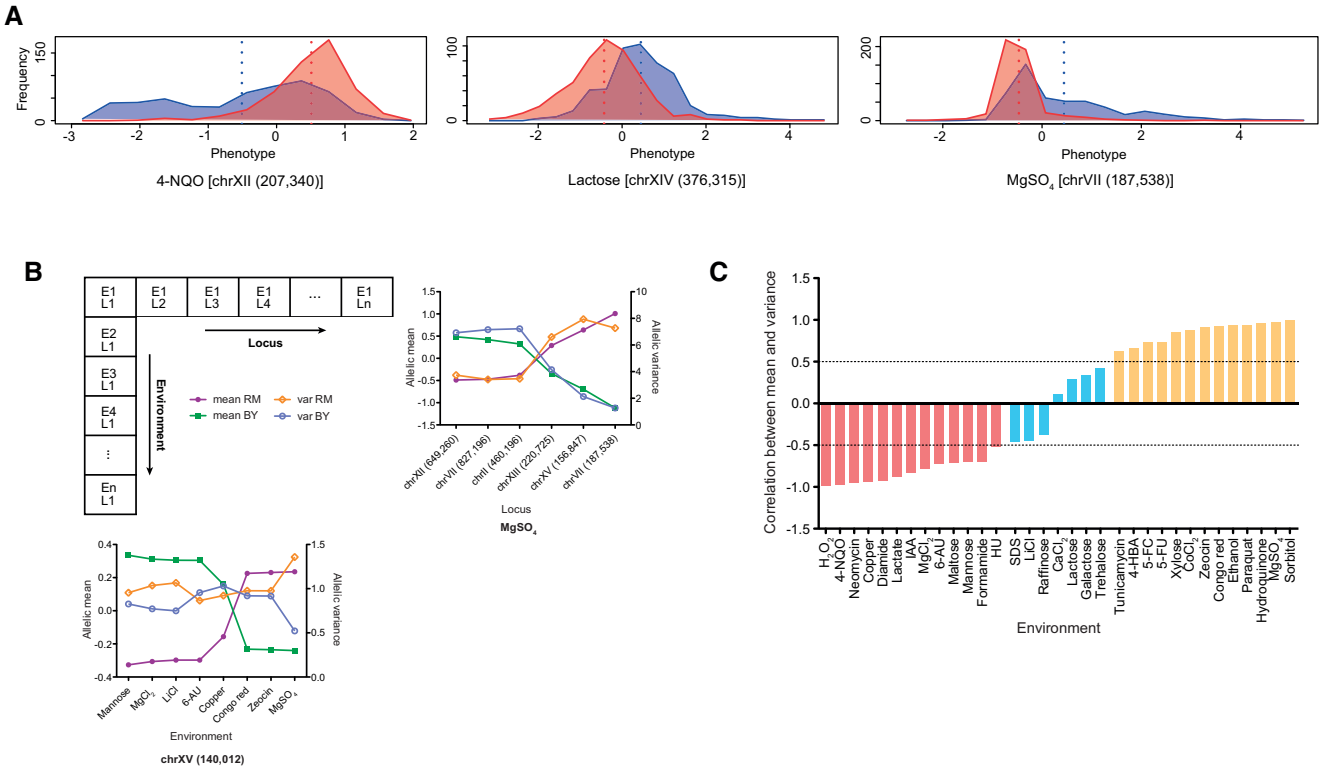

Fig. 2.—Environment and locus dependence of variance. (*A*) Representative frequency distributions of three loci showing directionality of release of variance. Blue distribution is of segregants with BY allele and red is for segregants with RM allele. 4-NQO [chrXII (207,340)] locus shows release of variance of the BY allele in the negative direction (mean RM > mean BY); Lactose [chrXIV (376,315)] locus shows equal variance of the two alleles; $MgSO_4$ [chrVII (187,538)] locus shows release of variance of BY allele in the positive direction (mean RM < mean BY). The loci are indicated as chromosome number followed by locus position in bp within brackets. (*B*) Environment and locus dependence of the mean and the variance of loci. The horizontal arm of the L-shaped grid represents multiple loci L_1–L_n significant in environment E1. The corresponding graph shows the mean and the variance of BY and RM alleles of all loci (*x*-axis) in $MgSO_4$. The vertical grid represents pleiotropic effect of the locus L1 in multiple environments E_1–E_n and corresponding graph shows mean and variance of the locus chrXV (140,012) in various environments (*x*-axis). (*C*) Correlation between the mean and the variance of all QTL in each environment. Correlations, $r^2 < -0.5$ (red; $P < 0.05$) represent environments that have a negative association between the mean and the variance. Correlations, $r^2 > 0.5$ (yellow; $P < 0.05$) represent environments that have a positive association between the mean and the variance. Environments with no directional release of variance (correlations, $-0.5 < r^2 < 0.5$) are represented as blue. The *y*-axis is correlation between mean and variance; dashed lines mark ±0.5 correlation value.

and the variance of alleles of loci within each environment (fig. 2*B*). Pearson correlation was calculated between the mean and the variance of alleles of all significant loci for each environment independently. A poor correlation across the environments would indicate that the release of variance could occur in either a positive or a negative mean direction, whereas a strong correlation would mean that release of variance occurred in only a specific direction (either positive or negative but never both) within an environment. Interestingly, a significant correlation ($P < 0.01$) was observed in majority of the environments (26/33, fig. 2*C*). Half of the environments (13/26) showed a strong positive correlation ($r^2 > 0.5$; Pearson correlation, $P < 0.01$) indicating that in these environments high variance was associated with higher phenotypic mean. The other half of environments showed a negative correlation ($r^2 < -0.5$; Pearson correlation, $P < 0.01$), i.e., high variance was associated with poor phenotypic mean (fig. 2*C*). For example, a strong positive correlation ($r^2 = 0.9$) was observed for all loci in $MgSO_4$ such that, independent of the locus, an allele with a higher mean always had a higher variance (fig. 2*B*). Even though in different environments, segregants showing both high and low phenotype contribute to high variance demonstrating that differential variance is not a result of a technical bias in phenotypic estimation but is a biologically relevant phenomenon. This meant that independent of the molecular nature of the regulatory locus and the genetic variation present in the population, the environment was the primary determinant of the effect of increase in variance on the population mean.

Pleiotropy of Loci Affecting Variance Underlies Gene-Environment Interactions

This prepotency of the environment in determining the direction of release of variance could potentially be the underlying basis of gene–environment interactions. In a previous study,

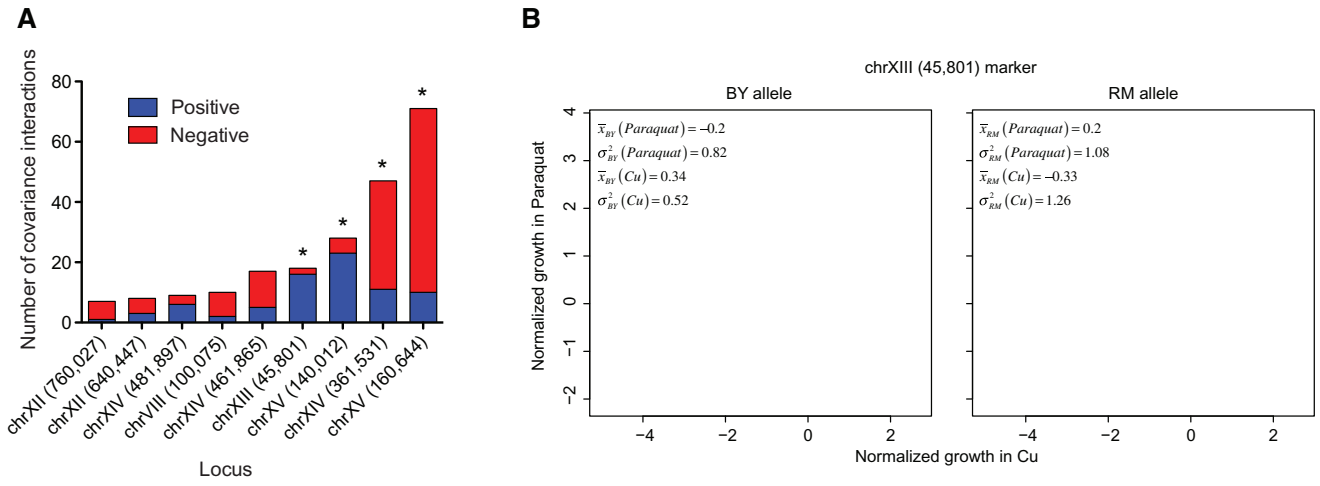

Fig. 3.—Allelic bias in across environment covariance. (*A*) Frequency distribution of the number of positive and negative values of Deming regression *t*-test (Methods) across pleiotropic covariance loci (*x*-axis). The *t*-test value is positive where Cov (BY) > Cov (RM) and negative for Cov (BY) < Cov (RM). The *y*-axis indicates the number of environmental pairs across which the locus showed a significant difference in covariance. An asterisk denotes a significant Fisher Exact test ($P < 0.05$). The locus is indicated as a chromosome number followed by the locus position in bp within brackets. (*B*) Covariance of normalized growth phenotype of the BY (blue) and the RM (red) allele of chrXIII (45,801) locus in Paraquat and Cu (Copper). The mean and the variance of each allele in each environment are indicated in the box. Paraquat shows a positive correlation between the mean and the variance whereas Cu shows a negative correlation (fig. 2*C*). The locus is indicated as a chromosome number followed by the locus position in bp within brackets.

we reported high antagonism of the mean effect of loci across multiple environments (Yadav et al. 2015). Thus, we next asked how the environment, by determining direction of release of variance, impacts the mean effects of pleiotropic loci. A locus can show environment dependent effect on variance in two ways; either it can result in difference in variance in one environment but not the other, or the identity of the segregants that show extreme phenotypes contributing to high variance can be different in the two environments. Therefore, in order to compare the effect of loci on variance across multiple environments in a comprehensive manner, we compared trait covariance of alleles across pairs of environments. Trait covariance measures how a population behaves across two environments (Haber and Dworkin 2016). A high covariance means similar phenotype of the segregants across a pair of environments, whereas a low covariance indicates uncorrelated phenotype of the segregants across the environments. As an alternative to performing a genome-wide trait covariance analysis, we chose only those loci (mQTL, vQTL, or mvQTL) that had a significant effect in more than one environment. Forty-two such loci were selected and the trait covariance of their alleles was compared across all possible pair-wise combinations of environments (Methods), where the locus had a significant effect (supplementary table S3, Supplementary Material online). Eighteen loci showed a significant difference (F-statistic, corrected $P < 0.05$) in the allelic covariance across at least one environmental pair, of which 5 loci were significant across more than 15 environmental pairs (fig. 3*A*, supplementary table S3, Supplementary Material online). We observed that low covariance was often a property of the allele that had higher variance between the two environments compared. Further, we observed that the loci with significantly different covariance across multiple pairs of environments showed an allelic bias. Among the 5 loci that showed differential covariance in more than 15 environmental pairs, 4 had a significant allelic bias, i.e., one allele had a tendency to show higher covariance than the other allele (Fisher Exact test, $P < 0.05$, fig. 3*A*). This was independent of the association between mean and variance in the environments compared. For example, in fig. 3*B*, Paraquat shows a positive correlation between the mean and the variance whereas Copper (Cu) shows a negative correlation (fig. 2*C*). However, BY allele of the chrXIII (45,801) locus shows reduced variance in both environments whereas the RM allele shows a higher variance in both. Hence, independent of the environment, alleles of a locus tend to either increase variance or decrease it across most environments in a consistent manner.

This consistent effect of alleles on covariance, along with strong environmental control of mean and variance, has profound implications on the nature of gene–environment interactions across diverse environments. Assume a following scenario where a locus shows differential covariance in two environments that show strong but opposite correlation between the mean and the variance. Because there is a high consistency in release of variance, one allele will result in high variance in both environments, whereas the other allele will show lower variance. The opposite correlation of mean and variance in the two environments would result in opposite effects on the allele with high variance in the two environments resulting in antagonistic pleiotropy. As a case in point,

in figure 3*B*, BY allele of chrXIII (45,801) locus shows antagonism for mean values between Cu and Paraquat, the underlying basis of which is high variance of RM allele in both environments but differential effect of high variance in two environments.

Difference in Variance is an Accurate Predictor of Genetic Interactions

The genetic interpretation of differential variance is that an allele with higher variance allows phenotypic manifestation of other genetic variants in the population, whereas the allele with lower variance buffers their effects. This implies that the vQTL and mvQTL regulate the effect of other loci, i.e., show genetic interactions. If this inference is correct then we should observe an increased tendency of these loci to show two-locus interactions. Additionally, just like a single locus can affect mean, variance or both, two loci can also interact to affect mean, variance or both. We carried out conventional two-locus mapping (2-mQTL mapping) that identifies two loci which interact to affect phenotypic mean. In addition, we modified the two-locus mapping technique using BF-statistic to identify genetic interactions that affect the variance of the segregants (2-vQTL mapping). If interaction between two loci resulted in a difference in both mean and variance, we termed it as 2-mvQTL interaction.

A 2-mQTL and 2-vQTL mapping was carried out in each environment independently between all loci that were significant as mQTL, vQTL, or mvQTL in those environments. A total of 73 significant interactions ($P < 0.1$) were identified of which 18 (25%) were 2-mQTL, 33 (45%) were 2-vQTL and 22 (30%) were 2-mvQTL interactions (fig. 1*C*). This substantially high number of 2-vQTL interactions, in comparison to a relatively small proportion of single vQTL (fig. 1*B*), highlights the role of genetic interactions contributing to population variation without demonstrating epistasis of the mean effects. This may explain the lack of identification of two-locus interactions in other studies that focus only on mean effects. Whereas we did not identify more number of loci compared with Bloom et al. (2013), using 2-vQTL mapping we identified a substantially high number of two-locus interactions than previously observed. Most of the 2-mQTL and 2-mvQTL interactions were identified in environments that showed a large difference between broad (H^2) and narrow (h^2) sense heritabilities, i.e., the high heritability was a result of epistatic interactions instead of primarily additive effects. However, Bloom et al. (2013) did not identify two-locus interactions in all of these environments. For example, for Hydroquinone, broad sense heritability (H^2) was 0.6, but the additive QTL identified could only explain a variance of 0.24 in the population and did not identify any two-locus interactions (Bloom et al. 2013). Our 2-vQTL mapping approach identified three 2-vQTL interactions that could potentially explain this difference between broad and narrow sense heritabilities observed in this environment (supplementary table S4, Supplementary Material online). Interestingly, amongst the loci showing 2-mQTL interactions, only 10% had mQTL effects, whereas 80% were either vQTL or mvQTL. Moreover, a negative correlation was observed between the total number of mQTL and the numbers of 2-mQTL ($r^2 = -0.46$) and 2-vQTL ($r^2 = -0.37$) interactions across environments. On the other hand, a strongly positive correlation was observed between number of mvQTL and numbers of 2-mQTL ($r^2 = 0.64$) and 2-vQTL ($r^2 = 0.65$) interactions. This further supports our previous observation of two types of genetic architectures, one consisting predominantly of mQTL and other of mvQTL with higher abundance of epistatic interactions in environments constituting mainly of mvQTL. Along with demonstrating a genetic basis of difference in variance, we show that difference in phenotypic variance is a robust predictor of tendency of a locus to show genetic interactions.

A higher likelihood of mvQTL to show genetic interactions provides an evidence for a genetic basis of why one allele of vQTL has higher variance compared with the other allele. The differential variance and covariance of the two alleles is a due to the phenotypic manifestation of other genetic loci, or CGV, in the presence of one allele but not the other. At a genetic level, it indicates that CGV is released in the presence of one allele (with higher variance) and buffered in the presence of the other (with lower variance). As an example to explain the above argument, chrXIV (368,185) locus shows differential allelic covariance across 4-HBA and Galactose and has distinct genetic interactions in both the environments. The BY allele of the chrXIV (368,185) locus has a higher variance in both 4-HBA and Galactose whereas RM allele has a lower variance (fig. 4*A*). In 4-HBA, the chrXIII locus shows a larger effect in the presence of the BY allele of chrXIV locus than the RM allele (fig. 4*B*). Similarly, in Galactose, chrXV locus shows a stronger effect on the phenotype in the presence of BY allele of chrXIV locus than the RM allele (fig. 4*C*). Similar results were also observed for other genetic interactions as well (supplementary table S4, Supplementary Material online).

Environment Specific Combinations of Genetic Interactions Regulate Phenotypic Variance

Enrichment of mvQTL in a few environments (fig. 1*D*) and higher tendency of mvQTL to show genetic interactions, these two observations together indicate that whereas genetic interactions are abundant in quantitative traits, they are not uniformly distributed. In other words, for the same population and the same phenotype (colony size), some environments reveal multiple genetic interactions whereas others show a primarily additive genetic regulation. Additionally, whereas these genetic interactions are environment specific, we observed that the same loci are involved in multiple genetic interactions within an environment (fig. 5*A*). Furthermore, most of the interactions with overlapping interacting loci were either 2-vQTL or 2-mvQTL, i.e., these loci affected variance

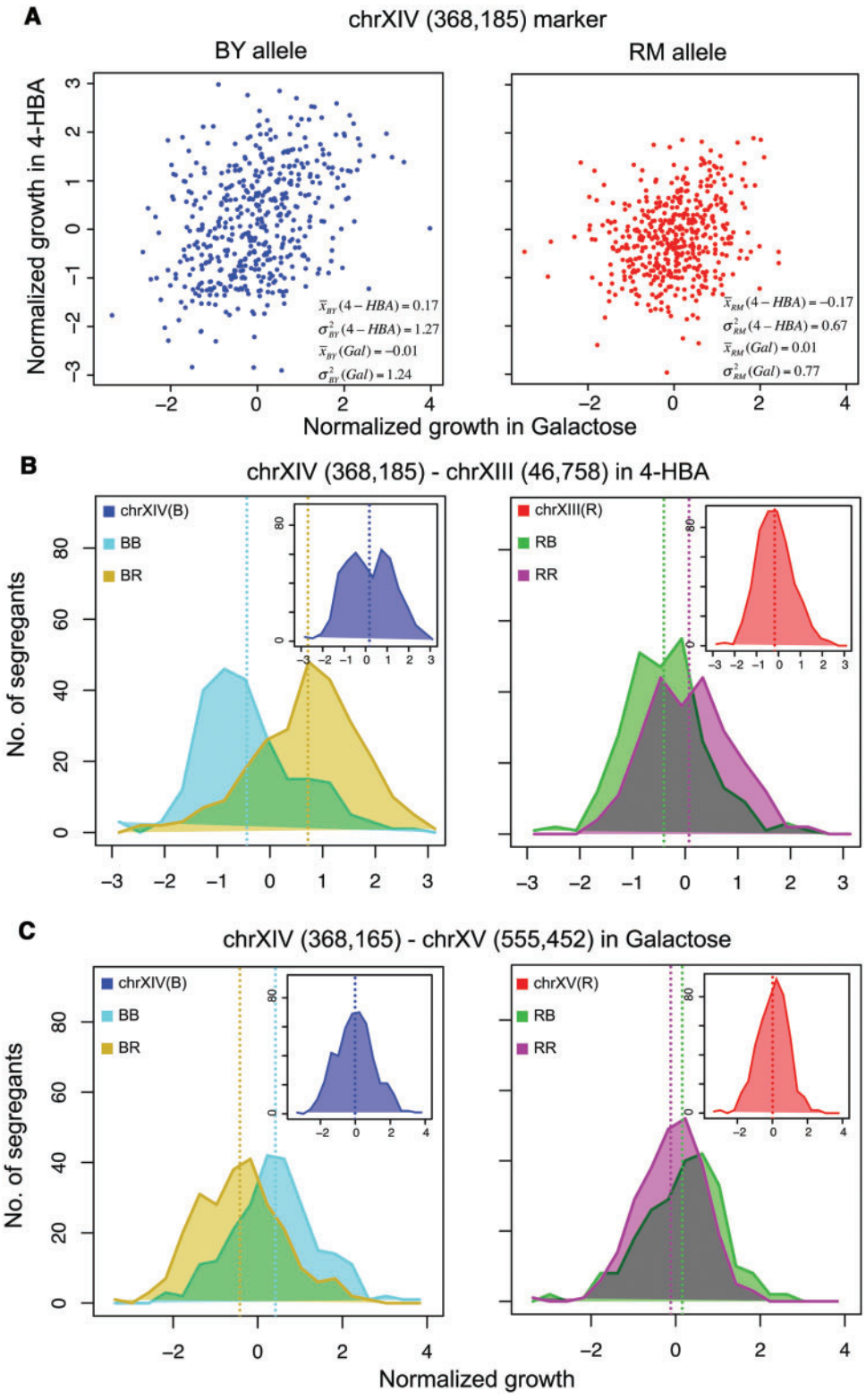

Fig. 4.—Two-locus interactions showing phenotypic manifestation of CGV. (*A*) Covariance of normalized growth phenotype of the BY (blue) and the RM (red) segregants for chrXIV (368,185) locus in 4-HBA and Galactose. The mean and the variance of each allele in each environment are indicated in the box. The axes are normalized growth of segregants in the two environments indicated. (*B*) Two-locus interaction between chrXIV (368,185)–chrXIII (46,758) in 4-HBA. (*C*) Two-locus interaction between chrXIV (368,165) - chrXV (555,452) in Galactose. The locus is indicated as a chromosome number followed by

of the segregating population in a redundant manner within an environment. As a case in point, the four interacting loci identified in Indoleacetic acid (IAA) show significant 2-mvQTL interactions with each other (fig. 5*B* and *C*). In IAA, a negative correlation (fig. 2*C*) was observed between mean and variance such that alleles with higher mean are associated with low variance. In each two-locus interaction, only 1 of 4 biallelic combinations has higher variance and lower mean phenotype compared with other combinations (fig. 5*B*). A significant two-locus interaction between all six pairs of these loci is indicative of existence of a higher order interaction among them. While we did not perform multi-locus interaction mapping, this implies presence of a higher order genetic interaction network that regulates phenotypic variance. Most allelic combinations of this network result in low variance, i.e., buffered phenotype, but some allelic combinations result in release of variance. Similar such genetic networks were identified in other environments based on comparison of two-locus interactions (fig. 6). As commented previously, these networks were identified in environments with high broad but low narrow sense heritabilities, indicating a high contribution of epistatic interactions in their regulation, e.g., IAA and $MgSO_4$ have H^2 of 0.81 and 0.87 but h^2 of only 0.44 and 0.43, respectively. Whereas the degree of interactivity of a locus varied within each environment, all interactions affected just variance or both mean and variance. Furthermore, we observed that these potential genetic interaction networks were highly environment specific. Even though pleiotropic loci were part of these interactions, the effect of these loci in such environments depended on the allelic combinations of other loci in the network. These results suggest a tempting possibility that a dense genetic interactome underlies all traits, which is revealed under very specific genetic combinations and environmental cues.

The loci involved in genetic networks are large effect loci affecting both mean and variance. To discover the candidate molecular network underlying the genetic networks identified in our study, we extracted the list of genes within these interacting loci using Bloom et al. (2013) data (supplementary table S5, Supplementary Material online). We used protein–protein interactions, experimental evidence and co-expression evidences in STRING database (Methods) to identify a network of interacting genes for each environment specific genetic network. We deduced candidate molecular networks for 5 of the 8 genetic networks (fig. 6) identified in our study, revealing potential environment specific molecular networks that maintain CGV and hence regulate the phenotype (supplementary fig. S1, Supplementary Material online). Interestingly, despite the high environment specificity of these networks, the candidate genes identified are associated with central cellular processes viz., stress response transcription factors, protein chaperones, ribosomal biogenesis, etc. (fig. 5*B* and *C*). This indicates that regulation of CGV may not be brought about by specialized set of genes but could be a fundamental property inherent to cellular functioning. Moreover, such an approach can be used to identify environment specificity of pleiotropic molecular regulators. As an example, a similar genetic architecture regulating phenotypic variation was identified in two different types of cellular stresses, IAA and Formamide. IAA is a plant hormone that stimulates adhesion mediated by Yap1 and Flo11 at lower concentrations and growth arrest at high concentration (Prusty et al. 2004; Spaepen and Vanderleyden 2011). On the other hand, Formamide acts as a stressor of biomolecules by affecting hydrogen bonds and invokes a general stress response (Hampsey 1997). Despite different known molecular responses, identical single QTL, covariance and genetic interactions (fig. 5*C* and *D*) were identified in both these environments. We identified a significant protein–protein interaction network of 10 genes amongst the 61 genes present in these specific genetic networks (PPI enrichment, $P = 4.7 \times 10^{-9}$; supplementary table S5, Supplementary Material online). This network of proteins is enriched in ribosome pathway (Bonferroni corrected FDR = 0.0073), specifically related to mitochondrial regulation. These mitochondrial regulators have not been previously associated with yeast growth in either IAA or Formamide. Additionally, that such constitutive ribosomal network could have potentially evolved as result of adaptation to specific selection pressures like IAA or Formamide. Hence, by overlaying molecular information over genetic networks regulating CGV, we have identified possible interactome that may have evolved in response to adaptation to diverse stresses.

Discussion

In this paper, by identifying loci that affect mean and variance in a diverse array of environments, we uncovered the regulation of phenotypic variation beyond a relatively simplistic additive architecture of genetic variation. We show that CGV is ingrained in the architecture of complex traits and understanding its regulation can uncover the genetic networks underlying the genotype–phenotype map. Our results uncover a hierarchical regulation of phenotypic manifestation of genetic variants where the large effect regulators buffer the effects of cryptic variants in most environments and the variants are

Fig. 4.—Continued

the locus position in bp within brackets. For (B, C), the *x*-axis is normalized growth of segregants in the environment and the *y*-axis of number of segregants. Dash lines in segregant distributions (B, C) indicate the means of the distributions. The biallelic locus segregant distributions (in the locus order written above the plots) are indicated as BB (light blue), BR (light brown), RB (dark green) and RR (purple). Inset plots show the average distributions of the first locus (BY (blue) and RM (red) alleles). See supplementary table S4, Supplementary Material online for details.

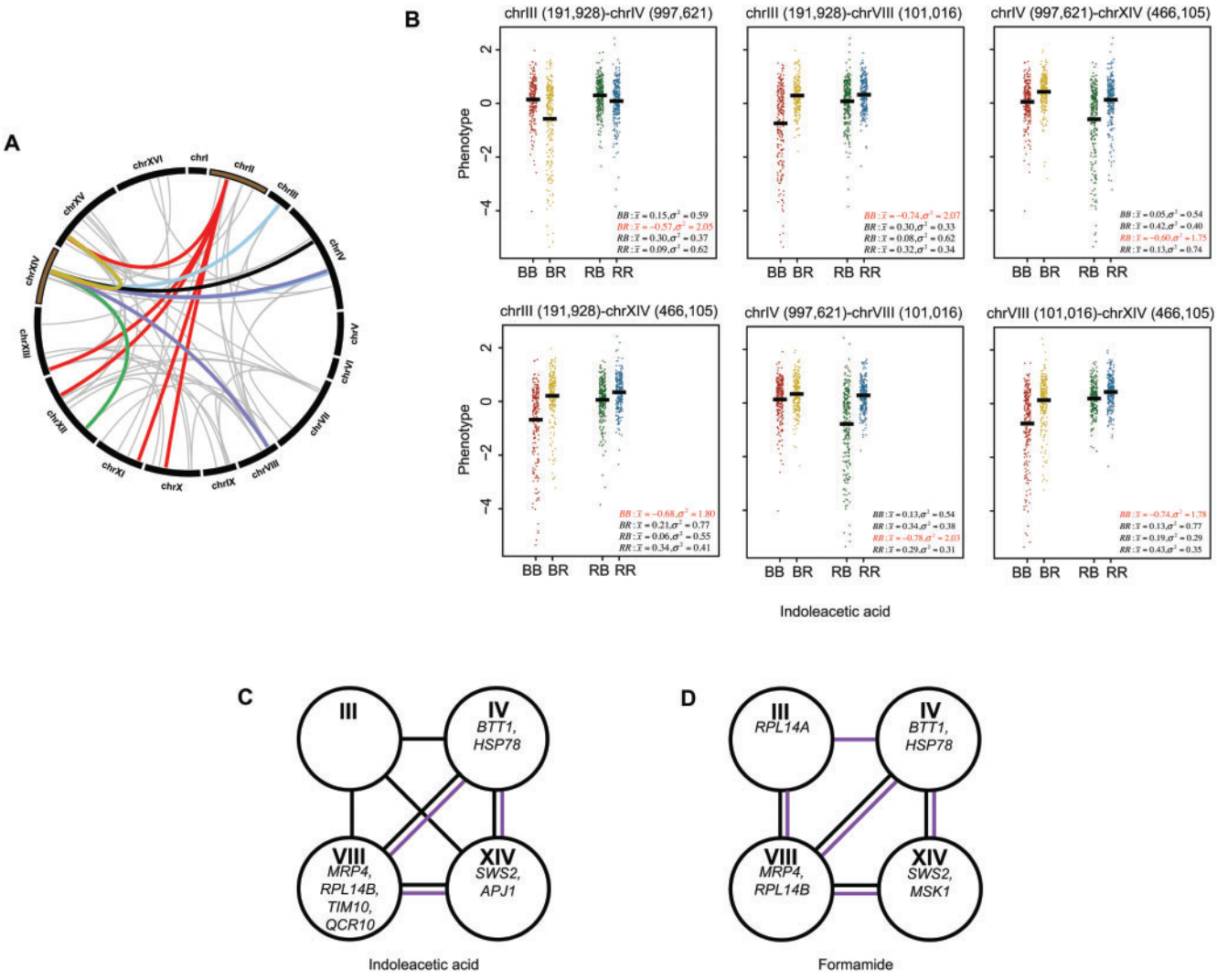

Fig. 5.—Two-locus interactions identify genetic networks. (*A*) Two-locus interactions between various loci shown as connected links. The chrII (245,879) locus (red) has multiple interactions (deep red) for growth in the same environment (Congo red). The chrXIV (466,590) locus (green) has 6 environment-specific two-locus interactions (4-NQO = dark green, Formamide = deep purple, Indoleacetic acid (IAA) = light blue, Lithium chloride (LiCl) = deep orange, Trehalose = deep yellow, Xylose = orange). Other two-locus interactions are indicated as light grey links. Genomic interaction maps made using Circos (Krzywinski et al. 2009). See supplementary table S4, Supplementary Material online for data. (*B*) Scatter plots showing examples of two-locus interactions of four loci [chrIII (191,928), chrIV (997,621), chrVIII (101,016), chrXIV (466,105)] in Indoleacetic acid (IAA). The biallelic locus segregant distributions (in the locus order written above the plots) are indicated as BB (red), BR (yellow), RB (green), and RR (blue) on *x*-axis. The mean and the variance of each allelic pair are indicated in the box with allelic pair with maximum value of variance indicated in red. The *y*-axis is normalized growth phenotype. (*C*) Schematic representation of genetic network of four loci, indicated in (B) above, showing two-locus interactions in Indoleacetic acid. Only the genes that showed molecular interactions are shown. The black lines indicate genetic interaction identified by two-locus analysis and purple line indicates molecular interactions identified through STRING (Methods). (*D*) Schematic representation of genetic network of four loci showing two-locus interactions in Formamide. Only the genes that showed molecular interactions are shown. The black lines indicate genetic interaction identified by two-locus analysis and purple line indicates molecular interactions identified through STRING.

revealed only upon a change in genetic background or under certain environments. These specific environments, probably determined by the adaptive history of the strains, perturb the internal buffering mechanisms that maintain the homeostasis between genetic and phenotypic diversity. This homeostasis is probably the main reason behind the primarily additive architecture of complex traits such that the underlying genetic networks are revealed only upon its perturbation. The genetic networks identified show two key properties. First, a single unique genetic network is active within each environment.

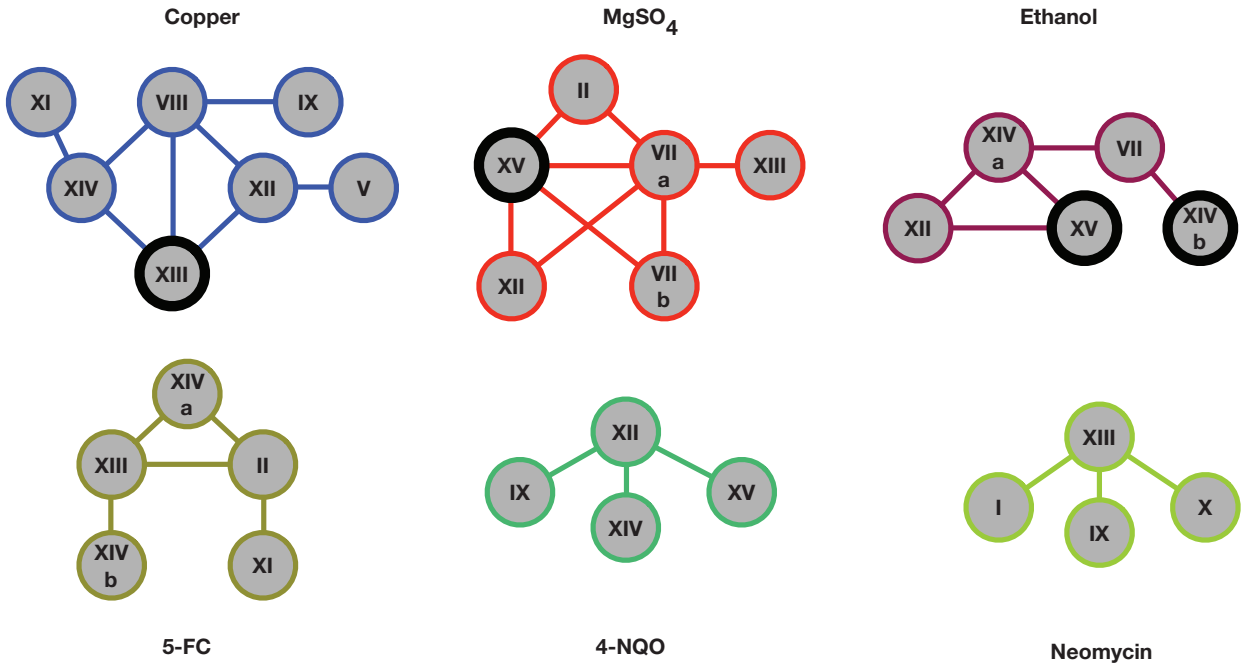

Fig. 6.—Genetic networks identified in diverse environments. (*A*) Examples of specific genetic networks identified in different environments using two-locus interaction mapping. Pleiotropic loci identified through covariance mapping (see fig. 3*B*) are highlighted as nodes with black outlines. See supplementary table S4, Supplementary Material online for details.

Even though some loci are pleiotropic across environments, the environment specific genetic networks determine their effects in each environment. Second, each environment and hence each network release CGV in only one direction, i.e., the released CGV either has a positive or a negative effect on mean but never both.

It is possible that some of the mvQTL identified by us can be composed of a vQTL linked to an mQTL. However, the mvQTL distribution observed in the study was highly skewed (fig. 1*D*) to be a result of such linkage. In some environments, most loci affected only mean (mQTL) without significant effect on phenotypic variance, whereas in other environments most loci affected both mean and variance (mvQTL), i.e., they behaved as both mQTL and vQTL. If an mvQTL was a result of linkage between an mQTL and a vQTL, then we would have observed some mvQTL in most environments, with the rest loci behaving as mQTL and vQTL. Such a clear demarcation of the presence of mvQTL, as observed in our study, rules out presence of linked loci affecting mean and variance independently as the major reason of mvQTL.

Gene–environment interactions (GEI) are a fundamental property of quantitative traits, such that all loci show environment dependence, albeit to different degrees (Mackay 2001). Whereas such an environmental dependence could be due to multiple reasons, our results indicate that differential regulation of CGV is a major contributor to GEI. Small effect GEI can arise if genetic variants are buffered in one environment such that they have no or small effects in that environments and revealed in the other environments thereby resulting in environment specific phenotypic effects. This explains the abundant small effect GEI observed previously (Mackay 2001; Yadav et al. 2015). This has implications towards understanding the environmental influence on phenotype, especially to explain the increasing rate of identification of trade-offs in various traits. In our previous paper (Yadav et al. 2015), we show that the environments like different carbon sources showed less trade-off among themselves than with other forms of stresses. In this paper, we show that a potential reason behind that could be buffered CGV in different carbon sources compared with stresses. Furthermore, we demonstrate how large scale GEI, especially antagonistic pleiotropy, could be due to differential association between mean and variance in various environments (fig. 3*B*). We believe that this can explain the increasing identification of antagonistic effects in various disease causing alleles (Carter and Nguyen 2011). Our study indicates that the trade-off of causal loci could be a result of differential buffering of CGV in different populations. This CGV is increasingly being uncovered due to population admixture and diverse lifestyles and can be a significant contributor to increasing incidence of modern complex diseases (Gibson and Reed 2008).

Environment is the primary determinant of the regulation of CGV such that it is released in only one direction in a particular environment. Even though other studies have focused on a few environments or phenotypes, we observed a similar

"directionality" of the released CGV in them, for example in studies interrogating the effects of Hsp90 (Jarosz et al. 2010) and other vQTL studies (Lee et al. 2014). Our multi-environment comparison shows that this is a property conserved across many environments. Because we show that single network of genetic loci is active in each environment, this directionality can be extrapolated to be a property of the genetic network. Multiple reasons can underlie this directional regulation of the effects of genetic variants. We propose that a genetic network, evolved for regulation of a phenotype, can work in two ways – the maximalistic or the minimalistic approach. The maximalistic approach would be evolution of networks that maintain the phenotype in a default "high" state. In this scenario, any genetic variant that results in a lower phenotype will not have substantial effect because an excess of "high" signal will mask its effect. As a result, such variants will get accumulated due to lack of purifying selection on them. Concurrently, a variant that increases the phenotype further will get fixed and become a part of the network. Therefore, perturbation in the buffering state of this network will result in phenotypic manifestation of only genetic variants that result in a poorer phenotype. The alternative, minimalistic approach would describe a default "low" state, such that the network maintains the phenotype in a nearly shut down state. With the upstream signal off or minimal, the downstream genetic variants that increase the flux will not be able to show their effects and hence get accumulated. Consequently, a breakdown in buffering of this network will reveal genetic variants that increase the phenotype. This regulation may happen at various levels (transcription, translation, or for metabolites) and can differ for different phenotypes and environments based on evolutionary histories of the populations. We would like to note that although weak and rare, a few loci affecting only variance without any significant effect on mean (vQTL) were also identified in our study. We have refrained from discussing properties of vQTL because their small number limits the ability to make any generalized observations. However, as an extrapolation from our two-QTL results, a vQTL would result in phenotypic manifestation of genetic variants in both directions and thus, not follow the two approaches as described above. It is therefore possible that different selection pressures would act on vQTL and mvQTL and future studies in this area will crucial to understand the differences in evolutionary importance of these two types of loci.

Finally, our results form a link between the genotype–phenotype map identified by studying population variation and genome-wide molecular studies. The environment specific genetic networks identified in our study reflect the trait and disease specific modular networks identified by various molecular studies. Integrating the two will help understand how the molecular networks are modified in diverse individuals and populations to regulate various phenotypes. Moreover, integrating the information from mapping studies and molecular studies can throw light upon the mechanistic basis of regulation of CGV and evolution of various networks.

Identification of molecular networks corresponding to genetic networks suggests that regulation of CGV is an integral part of adaptation and is potentially a major force shaping evolution of regulatory networks. We would like to highlight that identification of such genetic and molecular networks would not have been possible through conventional approaches for following reasons. First, these loci could not be identified in isogenic strains because a genetically heterogeneous population is required to detect differential effects on other genetic variants. Second, our study indicates existence of higher order genetic interactions that regulate CGV and hence, population phenotype. Identification of these multiallelic higher order interactions will difficult through deletion studies. Our results propose a high prevalence of such higher order interactions and emphasize the need to overlap genetic and molecular networks to comprehensively understand the genotype–phenotype map.

We propose that studying both mean and variance heterogeneity can address various long-standing questions in regulation of complex traits. The large effect loci identified for various phenotypes may not be directly causal but buffer the causal variants such that they show their effects only under certain allelic combinations of the genetic networks. This may explain the genetic background dependence and differential penetrance of causal loci that result in missing heritability in various GWA studies (Manolio et al. 2009; Eichler et al. 2010). Additionally, whereas molecular studies in model systems and cell lines have identified extensive networks associated with various diseases, we do not yet understand which of the sub-networks within these networks result in variation in disease predisposition in humans. This limits both disease prediction as well as identification of therapeutic targets within these large networks. Our study demonstrates that how studying variance heterogeneity and overlapping identified genetic networks with known molecular networks has the potential to address these limitations. The ease of yeast model allowed discovery of such genetic networks, therefore, revisiting the already existing datasets of human diseases in the light of conclusions from this study can provide better insights into mechanisms regulating variation in human diseases.

Supplementary Material

Supplementary tables S1–S5, figure S1, and file S1 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was supported by Tata Institute of Fundamental Research intramural funds (12P-0120 to H.S.). The funders

had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Literature Cited

Bhatia A, et al. 2014. Yeast growth plasticity is regulated by environment-specific multi-QTL interactions. G3 (Bethesda) 4:769–777.
Bloom JS, et al. 2015. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. Nat Commun. 6:8712.
Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. 2013. Finding the sources of missing heritability in a yeast cross. Nature 494:234–237.
Broman KW, Sen S. 2009. A Guide to QTL Mapping with R/qtl. Spinger: New York.
Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.
Camargo A, Azuaje F, Wang H, Zheng H. 2008. Permutation—based statistical tests for multiple hypotheses. Source Code Biol Med. 3:15.
Carter AJR, Nguyen AQ. 2011. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. BMC Med Genet. 12:160.
Costanzo M, et al. 2010. The genetic landscape of a cell. Science 327:425–431.
Dworkin I, Palsson A, Birdsall K, Gibson G. 2003. Evidence that *EGFR* contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila melanogaster*. Curr Biol. 13:1888–1893.
Eichler EE, et al. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 11:446–450.
Forsberg SKG, et al. 2015. The multi-allelic genetic architecture of a variance-heterogeneity locus for molybdenum concentration in leaves acts as a source of unexplained additive genetic variance. PLoS Genet. 11:e1005648.
Gibson G, Dworkin I. 2004. Uncovering cryptic genetic variation. Nat Rev Genet. 5:681–690.
Gibson G, Reed LK. 2008. Cryptic genetic variation. Curr Biol. 18:R989–R990.
Gibson G, Wemple M, van Helden S. 1999. Potential variance affecting homeotic *Ultrabithorax* and *Antennapedia* phenotypes in *Drosophila melanogaster*. Genetics 151:1081–1091.
Haber A, Dworkin I. 2016. Dis-integrating the fly: a mutational perspective on phenotypic integration and covariation. Evolution doi:10:1111/evo.13100.
Hampsey M. 1997. A review of phenotypes in *Saccharomyces cerevisiae*. Yeast 13:1099–1133.
Hermisson J, Wagner GP. 2004. The population genetic theory of hidden variation and genetic robustness. Genetics 168:2271–2284.
Jarosz DF, Taipale M, Lindquist S. 2010. Protein homeostasis and the phenotypic manifestation of genetic diversity: principles and mechanisms. Annu Rev Genet. 44:189–216.
Krzywinski MI, et al. 2009. Circos: An information aesthetic for comparative genomics. Genome Res. 19:1639–1645.
Lee C-R, Anderson JT, Mitchell-Olds T. 2014. Unifying genetic canalization, genetic constraint, and genotype-by-environment interaction: QTL by genomic background by environment interaction of flowering time in *Boechera stricta*. PLoS Genet. 10:e1004727.
Lempe J, Lachowiec J, Sullivan AM, Queitsch C. 2013. Molecular mechanisms of robustness in plants. Curr Opin Plant Biol. 16:62–69.
Liti G, et al. 2009. Population genomics of domestic and wild yeasts. Nature 458:337–341.
Mackay T. 2001. The genetic architecture of quantitative traits. Annu Rev Genet. 35:303–339.
Mackay TFC. 2014. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nat Rev Genet. 15:22–33.
Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. Nat Rev Genet. 10:565–577.
Manolio TA, et al. 2009. Finding the missing heritability of complex diseases. Nature 461:747–753.
Paré G, Cook NR, Ridker PM, Chasman DI. 2010. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the women's genome health study. PLoS Genet. 6:e1000981.
Phillips PC. 2008. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet. 9:855–867.
Prusty R, Grisafi P, Fink G. 2004. The plant hormone indoleacetic acid induces invasive growth in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 101:4153–4157.
Rönnegård L, Valdar W. 2011. Detecting major genetic loci controlling phenotypic variability in experimental crosses. Genetics 188:435–447.
Rönnegård L, Valdar W. 2012. Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. BMC Genet. 13:63.
Rutherford SL, Lindquist S. 1998. Hsp90 as a capacitor for morphological evolution. Nature 396:336–342.
Shen X, Pettersson M, Rönnegård L, Carlborg O. 2012. Inheritance beyond plain heritability: variance-controlling genes in *Arabidopsis thaliana*. PLoS Genet. 8:e1002839.
Siegal ML, Leu J-Y. 2014. On the nature and evolutionary impact of phenotypic robustness mechanisms. Annu Rev Ecol Evol Syst. 45:495–517.
Spaepen S, Vanderleyden J. 2011. Auxin and plant-microbe interactions. Cold Spring Harb Perspect Biol. 3:a001438–a001438.
Steinmetz LM, et al. 2002. Dissecting the architecture of a quantitative trait locus in yeast. Nature 416:326–330.
Szklarczyk D, et al. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43:D447–D452.
Taylor MB, Ehrenreich IM. 2015. Transcriptional derepression uncovers cryptic higher-order genetic interactions. PLoS Genet. 11:e1005606.
Vidal M, Cusick ME, Barabási A-L. 2011. Interactome networks and human disease. Cell 144:986–998.
Wu W-S, Lai F-J. 2015. Functional redundancy of transcription factors explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out. BMC Syst Biol. 9 Suppl 6:S2.
Yadav A, Radhakrishnan A, Bhanot G, Sinha H. 2015. Differential regulation of antagonistic pleiotropy in synthetic and natural populations suggests its role in adaptation. G3 (Bethesda) 5:699–709.

Associate editor: Gunter Wagner