UCLA
UCLA Previously Published Works

Title

Genetic interactions contribute less than additive effects to quantitative trait variation in yeast

Permalink

Journal

ISSN

Authors

Bloom, Joshua S
Kotenko, Iulia
Sadhu, Meru J
et al.

Publication Date

DOI

Genetic interactions contribute less than additive effects to quantitative trait variation in yeast

Joshua S. Bloom[1,2], Iulia Kotenko[3], Meru J. Sadhu[1], Sebastian Treusch[4], Frank W. Albert[1] & Leonid Kruglyak[1,2,5]

Genetic mapping studies of quantitative traits typically focus on detecting loci that contribute additively to trait variation. Genetic interactions are often proposed as a contributing factor to trait variation, but the relative contribution of interactions to trait variation is a subject of debate. Here we use a very large cross between two yeast strains to accurately estimate the fraction of phenotypic variance due to pairwise QTL–QTL interactions for 20 quantitative traits. We find that this fraction is 9% on average, substantially less than the contribution of additive QTL (43%). Statistically significant QTL–QTL pairs typically have small individual effect sizes, but collectively explain 40% of the pairwise interaction variance. We show that pairwise interaction variance is largely explained by pairs of loci at least one of which has a significant additive effect. These results refine our understanding of the genetic architecture of quantitative traits and help guide future mapping studies.

[1] Department of Human Genetics, University of California, Los Angeles, Los Angeles, California 90095, USA. [2] Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, California 90095, USA. [3] Department of Molecular Biology, Princeton University, Princeton, New Jersey 08540, USA. [4] Twist Bioscience, San Francisco, California 94158, USA. [5] Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, California 90095, USA. Correspondence and requests for materials should be addressed to L.K. (email: lkruglyak@mednet.ucla.edu).

Genetic interactions arise when the joint effect of alleles at two or more loci on a phenotype departs from simply adding up the effects of the alleles at each locus. Many examples of such interactions are known, but the relative contribution of interactions to trait variation is a subject of debate[1–5]. We previously generated a panel of 1,008 recombinant offspring ('segregants') from a cross between two strains of yeast: a widely used laboratory strain (BY) and an isolate from a vineyard (RM)[6]. Using this panel, we estimated the contribution of additive genetic factors to phenotypic variation (narrow-sense or additive heritability) for 46 traits and resolved nearly all of this contribution (on average 87%) to specific genome-wide significant quantitative trait loci (QTL). The repeatability of trait values across replicate measurements for each segregant provided an upper bound for the total contribution of genetic factors to phenotypic variation (broad-sense or full heritability). We used the difference between trait repeatability and the additive heritability as an estimate of the contribution of genetic interactions to trait variation. Because trait repeatability can include sources of variation other than gene–gene interactions, this approach can overestimate the contribution of such interactions. Further, with 1,008 segregants, we were able to detect only a small number of significant QTL–QTL interactions that, in aggregate, explained little of the estimated interaction variance.

Here we address these limitations by studying an expanded panel of 4,390 segregants obtained from the same cross. We genotyped these segregants at 28,820 unique variant sites and phenotyped them for 20 end-point growth traits with at least two replicates. The larger sample size permits us to directly and accurately quantify pairwise interaction variance, instead of relying on the difference between trait repeatability and the additive heritability. It also greatly increases the power to detect both additive QTL and QTL–QTL interactions (Supplementary Fig. 1). For example, we have 90% power to detect an additive QTL that explains 0.5% of phenotypic variance, and 90% power to detect a QTL–QTL interaction that explains of 0.8% of phenotypic variance (Methods section). Further, the expanded panel substantially improves fine mapping of loci.

We detected nearly 800 significant additive QTL. We were able to refine the location of the QTL explaining at least 1% of trait variance to ∼10 kb, and we resolved 31 QTL to single genes. We also detected over 200 significant QTL–QTL interactions; in most cases, one or both of the loci also had significant additive effects. For most traits studied, we detected one or a few additive QTL of large effect, plus many QTL and QTL–QTL interactions of small effect. We find that the contribution of QTL–QTL interactions to phenotypic variance is typically less than a quarter of the contribution of additive effects. These results provide a picture of the genetic contributions to quantitative traits at an unprecedented resolution.

Results

Partitioning trait variance. We used a linear mixed model with additive, pairwise interaction, and residual strain repeatability terms to quantify these components of trait variation[7]. The additive and interaction genetic contributions are estimated based on the realized relatedness[8,9] of all pairs of segregants, as measured from the dense genotype data. This approach allows us to separate the contribution of gene–gene interactions from other genetic and non-genetic sources of variation that can contribute to trait repeatability[7]. We used simulations (Methods section) to demonstrate that the model can accurately estimate the contributions of additive QTL and QTL–QTL interactions to trait variation over an extensive range of genetic architectures (Supplementary Fig. 2 and Supplementary Data 1).

Across the 20 traits, additive genetic variance ranged from 8.6 to 70.4% of phenotypic variance, with a median of 43.3%. Interaction genetic variance ranged from 2.2 to 21.2% of phenotypic variance, with a median of 9.2%. These measures provide genome-wide estimates for the aggregate effects of all additive and all pairwise interaction effects, respectively. The contribution of pairwise interactions to trait variance is typically less than a quarter of the contribution of additive effects, and does not exceed half the contribution of additive effects for any trait studied here. The remaining strain repeatability variance ranged from 0.05 to 21.4%, with a median of 8.8% (Fig. 1). Three-way interactions may account for some of the remaining effect of strain, but are unlikely to explain most of this remaining variance for most traits (Supplementary Data 2). This leaves higher-order interactions, other effects of strain, or experimental effects confounded with strain as the potential sources of the remaining strain repeatability variance.

Mapping additive QTL. Next, we sought to identify the individual genomic regions underlying these genome-wide estimates. We used a forward-search QTL mapping approach that controls for other QTL[10] (Methods section) to detect 797 genome-wide significant additive QTL, with a median of 42.5 per trait (range 17–56). We calculated the variance captured by these detected QTL with a random effect model that uses a genetic relationship matrix (GRM) constructed only from genotypes at the peak markers for each significant additive QTL. These loci captured a median of 92% of the additive genetic variance (Fig. 2a). The number of detected QTL per trait increased approximately fourfold relative to that in our previous study of a subset of 1,008 segregants from this panel[6], but the variance captured by significant QTL only increased by 5%, because most detected loci generally have very small effect sizes (median effect size of 0.38%; Fig. 4). These observations suggest that many additional undetected loci for these traits likely exist in this cross, but that their individual and collective effects are very small. The increased panel size also increases mapping resolution. The 180 loci that explain 1% or more of phenotypic variance have a median 95% confidence interval of 10.3 kb, compared with 31.2 kb with 1,008 segregants; these confidence intervals span ∼5 genes in the yeast genome. In 31 cases, QTL could be refined to a single gene (Supplementary Data 3).

Partitioning interaction variance. Detection of additive QTL that account for nearly all of the additive genetic variance allowed us to further partition the variance contributed by QTL–QTL interactions (Methods section). Briefly, we compared estimates of interaction variance captured by pairs of markers selected by three different criteria: all pairs of markers across the genome, the subset of pairs in which one marker is the peak of an additive QTL, and the subset where both markers are additive QTL peaks. As noted above, across the traits examined, the amount of phenotypic variance captured by interactions between all marker pairs had a median of 9.2%. The amount of phenotypic variance captured by interactions between significant additive QTL and the rest of the genome had approximately the same median (9.4%), whereas it dropped to 4.5% for interactions only between significant additive QTL (Fig. 3 and Supplementary Fig. 3). These results suggest that in most pairwise interactions, at least one of the loci has a significant additive effect, as can be confirmed by directly mapping QTL–QTL interactions (see below).

Mapping QTL–QTL interactions. We detected specific genome-wide significant QTL–QTL interactions for each trait using a statistically powerful approach that takes into account all the additive genetic variance (Methods section). One can test for
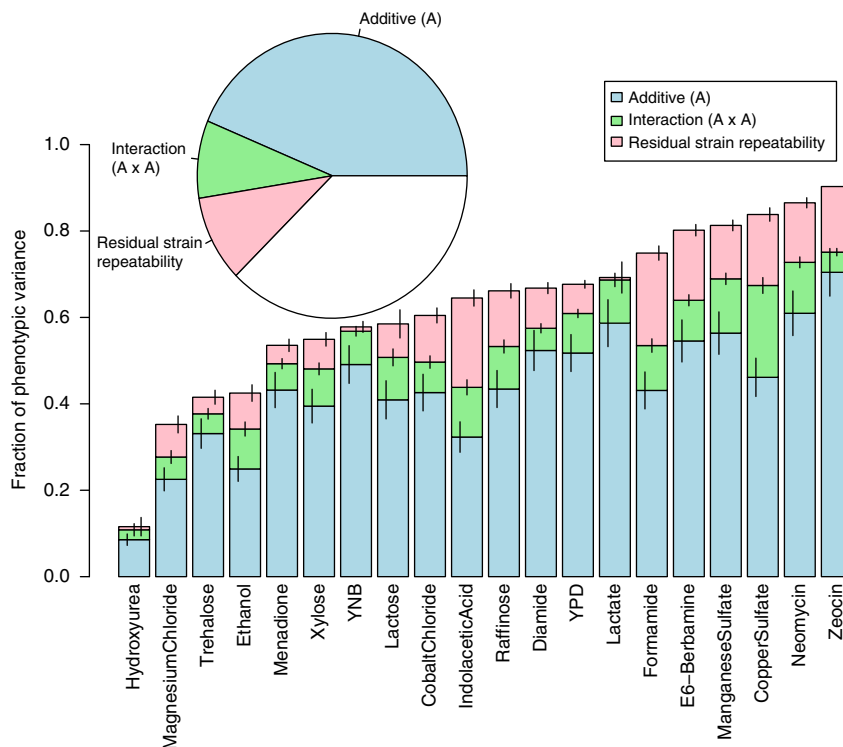
Figure 1 | Contributions to trait variation. Stacked bar plots of a variance component analysis for each trait are shown. The variance component model included terms for additive genetic variance (blue), two-way interaction variance (green), residual strain repeatability (pink) and residual error (not shown). Error bars show ± s.e. Inset, the average of the variance components across traits. Additive genetic effects, two-way interactions, and residual repeatability account for 43, 9 and 10% of phenotypic variance, respectively.
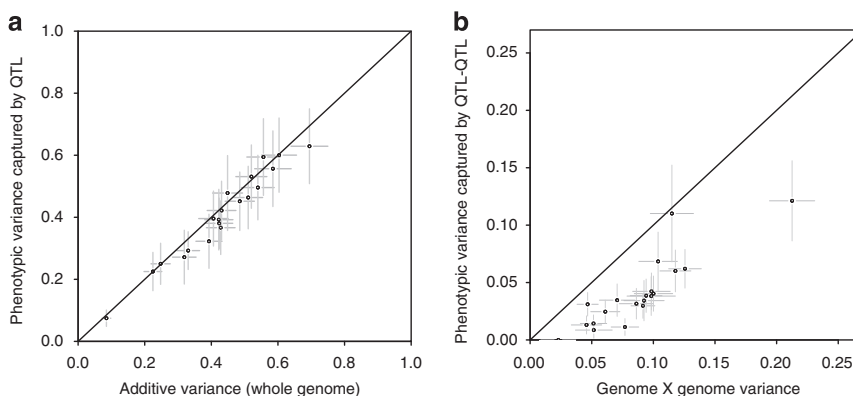

Figure 2 | Additive and interaction variance captured by detected loci. (**a**) Total variance captured by detected QTL for each trait is plotted against the whole-genome estimate of additive genetic variance. Error bars show ± s.e. The diagonal line represents (variance captured by detected QTL = additive genetic variance) and is shown as a visual guide. (**b**) Total variance captured by detected QTL–QTL interactions from the marginal scan for each trait is plotted against the whole-genome estimate of interaction variance. Error bars show ± s.e. The diagonal line represents (variance captured by detected QTL–QTL interactions = interaction genetic variance) and is shown as a visual guide.

interactions either between all pairs of markers (full scan), or only between pairs where one marker corresponds to a significant additive QTL (marginal scan). In principle, the former can detect a wider range of interactions, but the latter can have higher power due to a reduced search space. Here the two approaches yielded similar results, detecting 205 and 266 QTL–QTL interactions, respectively, at a false discovery rate (FDR) of 10%, with 172 interactions detected by both approaches. In the full scan, 153 of the QTL–QTL interactions correspond to cases where both interacting loci are also significant additive QTL, 36 correspond to cases where one of the loci is a significant additive QTL, and only 16 correspond to cases where neither locus is a significant

additive QTL (Supplementary Fig. 4 and Supplementary Data 4). The interactions detected in the full and marginal scans captured a median of 3.2 and 3.4% of phenotypic variance, respectively (Fig. 3). These numbers correspond to about 40% of the total pairwise interaction variance estimates (Fig. 2b), and greatly exceed expectations from background linkage effects[11] (Supplementary Fig. 5). Like the detected additive QTL, the detected QTL–QTL interactions generally have very small effect sizes, with a median variance explained of 0.31%. The remainder of the interaction variance is likely due to many more pairs with even smaller effect sizes. Unlike the case for additive QTL, no large-effect QTL–QTL interactions were observed for these 20
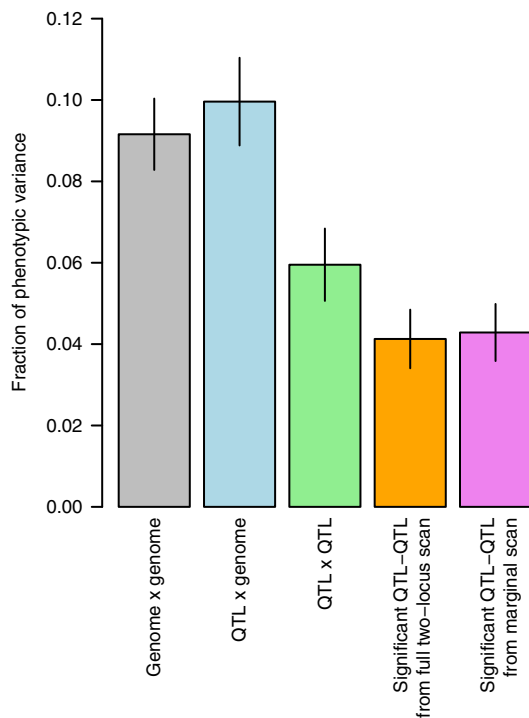
Figure 3 | Phenotype variance captured by different variance component models of two-way interactions. The average fraction of phenotypic variance captured by different variance component models of two-way interactions across traits. The bar heights represent variance estimated with all markers (genome × genome; grey), significant additive QTL by all markers (QTL × genome; blue), additive QTL by additive QTL (QTL × QTL; green), significant QTL–QTL detected from the marginal scan (orange), and significant QTL–QTL from the exhaustive two-dimensional scan (purple). Error bars show ± s.e.
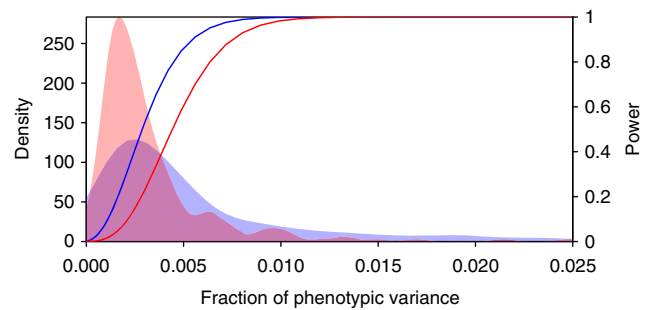

Figure 4 | Distribution of genetic effects and power to detect them. A density plot of the fraction of phenotypic variance (x axis) explained by individual significant QTL (blue area) and QTL–QTL interactions (red area) across all traits. The curves correspond to the statistical power at a genome-significance threshold (right, y axis) for QTL (blue) and QTL–QTL interactions (red).

traits. Whereas the largest additive QTL explained 26% of phenotypic variance, and 46 QTL had effect sizes > 5%, the largest QTL–QTL interaction explained only 3.3% of phenotypic variance (Fig. 4). Typical genetic architectures for traits in this study consist of a few large additive QTL and many small QTL and QTL–QTL interactions.

Discussion
We have used a very large yeast cross with 4,390 segregants to study quantitative trait variation in greater detail. Across 20 traits, we find that additive genetic effects and pairwise genetic interactions explain 43.7 and 9.2% of phenotypic variance, respectively, in agreement with previous estimates based on a smaller data set[12]. We detected a median of 42.5 significant additive QTL per trait. On average, these QTL captured 92% of the estimated additive heritability. Loci that explain at least 1% of phenotypic variance of loci typically spanned no more than 10 kb. We further estimate that roughly half of the pairwise interaction variance is contributed by interactions among significant additive QTL, and that nearly all of the interaction variance is contributed by interactions between significant additive QTL and the genome. Two-locus interactions in which neither locus has an additive effect are rare and do not contribute much to phenotypic variance. We detected about 13 QTL–QTL interactions per trait; these capture 3.2% of phenotypic variance or 40% of total pairwise interaction variance.

We previously discussed the factors that may lead to greater 'missing'[13] heritability in human genome-wide association studies

than in a yeast cross[6]. These include greater genetic variation captured by population studies, differences in the allele frequency spectrum, larger mutational target size of the human genome, higher physiological complexity, and within-locus dominance effects. Here we have focused on better delineating the contributions of pairwise interactions to phenotypic variance. The larger cross enabled us to obtain an accurate genome-wide estimate of these contributions, and revealed that they are substantially smaller than those of additive effects for every trait examined. Further, few interactions arise from locus pairs without detectable additive effects. This is consistent with what has been observed in reverse-genetic screens with gene knockouts[14]. Although accurate estimates of the contributions of higher-order interactions require even larger sample sizes, the preliminary estimates obtained here (Supplementary Data 2) suggest that such interactions contribute less than pairwise ones. Theoretical results have been used to argue that the contributions of interactions to phenotypic variance in outbred populations are expected to be smaller than in a cross[1,2]. We note that a small contribution of genetic interactions to trait variance does not imply that interactions do not exist, that they are not important for understanding the complete genetic basis of specific traits, or that genes do not act epistatically at the molecular level[5,14]. Individual examples of QTL–QTL interactions, including some of large effect, have been detected for a broad range of traits in many species[5,15–17]. In studies that have estimated the contribution of pairwise interactions to trait variance, it is often within the range observed here[18–21]. Our results further support the predominance of additive factors in explaining quantitative trait variance. They also suggest that interactions are most effectively detected by starting with the set of loci with additive effects. Combined with the recent observation of a small contribution of dominance to human trait variation[22], this suggests that heritability not captured by genome-wide additive models arises primarily from additive effects of variants untagged by current genotyping technologies[23].

Methods
Construction of segregant panel and sequencing libraries. The BYxRM segregants were constructed as described previously[6]. Before, we chose one segregant each from a panel of 1,184 dissected tetrads, ultimately analysing a panel of 1,008 segregants. Here we added segregants from this panel of tetrads that were not previously genotyped to assemble a new panel of 4,390 segregants. A Biomek FX liquid handling robot (Beckman Coulter) was used to re-array segregants that had not been previously genotyped to 1 ml of yeast peptone dextrose in 2-ml deep-well 96-well plates (Thermo Scientific). Plates were sealed with Breathe-Easy gas-permeable membranes (Sigma-Aldrich), and the yeasts were grown for 2 days at

30 °C without shaking. DNA was extracted using 96-well DNeasy Blood & Tissue kits (Qiagen). DNA concentrations were determined using the Quant-iT dsDNA High-Sensitivity DNA quantification kit (Invitrogen) and the Bio-Tek Synergy 2 plate reader. DNA was diluted to 0.2 ng µl^{-1}. Per sample, 5 µl of 0.2 ng µl^{-1} DNA was added to 4 µl of 5X Nextera HWM buffer (Illumina), 6 µl of water and 5 µl of 1/35 diluted Nextera enzyme. The transposition reaction was performed for 5 min at 55 °C. Illumina sequencing adaptors and custom indices were added by PCR, directly after the tagmentation reaction without additional sample purification. Fragmented DNA (10 µl) was combined with 0.5 µl each of 10 µM index primers (one of N701-N712 plus one of 96 custom indices), 5 µl of 10X Ex Taq buffer, 0.375 µl Ex Taq polymerase (Takara), 4 µl of 2.5 mM dNTPs and 29.625 µl of water, and amplified with 20 cycles of PCR. 1152-plex libraries were run on two single-end lanes of a rapid-run flow cell of a HiSeq 2500 (Illumina).

Power calculations. We calculated statistical power (1-β) for sample sizes of 100, 1,000 and 4,000 segregants in R using the 'power.t.test' function[24]. Power was calculated over a range of effect sizes, where effect size was calculated as the per cent phenotypic variance explained by a single QTL or QTL–QTL interaction. To correct for multiple testing genome-wide significance thresholds (α) of $P < 6.9 \times 10^{-4}$ and $P < 2.5 \times 10^{-5}$ were used for additive and interacting QTL, respectively. These thresholds were chosen based on a familywise error rate (FWER) <5% for the additive scan and FDR<10% for the interaction scan. We note that we used a less stringent 10% FDR threshold for detecting individual QTL–QTL interactions to provide greater sensitivity to detect interaction effects and we expect that detected interactions are likely more upwardly biased than additive QTL effect sizes.

Determining segregant genotypes. Fastq files for the 3,552 segregants sequenced for the present study were demultiplexed using fastq-multx[25] and aligned to the SacCer3 version of the reference genome using bwa (ref. 26). The 3,552 new segregants were sequenced with an average coverage of ~2X. The 1,056 previously sequenced segregants were realigned to SacCer3. BAM (ref. 27) files for all 4,608 segregants were merged into one BAM file and variants were called as described previously. An additional filter was used to remove regions with strong mapping bias towards the reference genome[28]. Of 39,741 high confidence single nucleotide polymorphisms at which BY and RM differ, 28,220 unique single nucleotide polymorphisms were retained for downstream analysis. As described previously, a hidden Markov model was used to infer the segregant genotypes[6]. Segregants were removed if they had fewer than 25 or greater than 105 recombination breakpoints, fewer than 35,000 markers with genotype calls, or if the segregant genotype was correlated with another segregant with a Pearson correlation >0.9. In all, 4,390 segregants passed these filters and were used for mapping.

Segregant phenotyping. All 4,390 segregants were phenotyped together, including the 1,056 previously characterized and sequenced segregants. Phenotyping was performed as described previously[6]. Briefly, segregants were pinned to agar plates from liquid stocks and then imaged for end-point growth at 48 h. Colony radii were calculated using functions in the EBImage R package[29]. End-point growth measurements were filtered and normalized as previously described. Traits with larger difference between broad- and narrow-sense heritabilities in our previous paper were prioritized here to focus on those traits more likely to have an appreciable contribution from genetic interactions. Therefore, the fraction of variance explained by genetic interactions could be biased upwards relative to all traits.

Segregant genotypes and phenotypes are available as Supplementary Data 5.

Calculating variance components. Custom R code was used to estimate variance components and map additive QTL as well as QTL–QTL interactions. A repeated measures mixed model[7] was used to estimate variance components. The model can be written as:

$$y = \beta X + Za + Zi + Zp + e$$

where y is a vector of length m that contains phenotypes for n segregants including replicate measurements such that $m = n*$ (number of replicates). β is a vector of estimated fixed effect coefficients. X is a matrix of fixed effects (here β is the overall mean, and X is a 1_m vector of ones unless otherwise specified). Z is an $m \times n$ incidence matrix that maps m total measures to n segregants. a is the additive genetic effects, i is the pairwise genetic interaction effects and p is the effects due to residual strain repeatability. The residual error is denoted by e. The distributions of these effects are assumed to be normal with mean zero and variance–covariance as follows:

$$a \sim N(0, \sigma_A^2 A), \quad i \sim N(0, \sigma_{AA}^2 A \circ A), \quad p \sim N(0, \sigma_R^2 I_n), \quad \text{and} \quad e \sim N(0, \sigma_{EV}^2 I_m)$$

The variance structure of the phenotypes is $V = \sigma_A^2 ZAZ' + \sigma_{AA}^2 ZA \circ AZ' + \sigma_R^2 ZI_n Z' + \sigma_{EV}^2 I_m$. Here, A is the additive relatedness matrix, the fraction of genome shared between pairs of segregants. A was calculated using the 'A.mat' function in the rrBLUP R package[30]. σ_A^2 is the additive genetic variance captured by markers. $A \circ A$ is the Hadamard (entrywise) product of A, which can be interpreted as the fraction of pairs of markers shared between pairs of segregants. σ_{AA}^2 is the

interaction genetic variance captured by all pairwise combinations of markers. I_n and I_m are $n \times n$ and $m \times m$ identity matrices, σ_R^2 is the residual effect of strain not captured by the additive and interaction genetic variance terms, and σ_{EV}^2 is the error variance. Variance components were estimated using AI-REML (ref. 31) and custom R code. Standard errors of variance component estimates were calculated as the square root of the diagonal of the Fisher information matrix from the iteration at convergence of the AI-REML algorithm.

An additional term for three-way interactions, using the Hadamard cube of A, is included in a model in Supplementary Data 2.

Mapping additive QTL. Additive QTL were mapped using a forward stepwise procedure. For each chromosome and trait the above model was fit, replacing the term for polygenic additive effects with a_{loco} where $a_{loco} \sim N(0, \sigma_{A\_loco}^2 A_{loco})$. $\sigma_{A\_loco}^2$ is the additive genetic variance from all chromosomes excluding the chromosome of interest, and A_{loco} is calculated as above, excluding markers from the target chromosome. The segregant best linear unbiased predictor (BLUP) residuals ($y_r = y - y_b$) for each chromosome were calculated by subtracting the BLUPs for the effects of the rest of the genome and pairwise interactions from the phenotypes, where $y_b = Z(\sigma_{A\_loco}^2 A_{loco} + \sigma_{AA}^2 A \circ A)Z'V^{-1}(y - X\beta)$ and $V = \sigma_{Am}^2 ZA_m Z' + \sigma_{AA}^2 ZA \circ AZ' + \sigma_R^2 ZI_n Z' + \sigma_{EV}^2 I_m$. Replicate values per strain were averaged. These averaged BLUP residuals for each chromosome were then used as the starting point for scans for additive QTL on the chromosome of interest. Using BLUP residuals increases power to detect QTL by controlling for genetic contributions from the remainder of the genome[10]. We tested for linkage at each marker on the given chromosome by calculating $(-n(\ln(1 - r^2)/2\ln(10)))$, where r is the Pearson correlation coefficient between the segregant genotypes at the marker and segregant BLUP residuals for n segregants. FWER thresholds were determined from empirical null distributions determined by recomputing the linkage statistic chromosome wide from 1,000 permutations of BLUP residual phenotypes to strain assignments and recording the maximum value[32]. The most significant marker was extracted from each QTL significant at a 5% FWER threshold. These peak markers were added to the model as fixed effects and residuals were recomputed. Additional linkage scans were performed on these residuals (using 5% FWER thresholds that were recomputed after each round of QTL addition) until no additional significant QTL were detected on that chromosome. Confidence intervals were calculated as 1.5 LOD drop using the lodint function in R/QTL (ref. 33).

Mapping QTL–QTL interactions. We increased power and computational efficiency by searching for interactions using the segregant BLUP residuals from the additive polygenic model as phenotypes. Specifically, we calculated y_r for each trait as $y_r = y - y_b$, where $y_b = Z(\sigma_A^2 A)Z'V^{-1}(y - X\beta)$ and $V = \sigma_A^2 ZAZ' + \sigma_R^2 ZI_n Z' + \sigma_{EV}^2 I_m$. Replicate values per strain were averaged. For the full two-dimensional scan, LOD scores for interactions were computed for all pairs of markers as $(-n(\ln(1 - r^2)/2\ln(10)))$, where n is the number of segregants with phenotypes, and r is the Pearson correlation coefficient between the product of segregant genotypes at pairs of markers separated by at least 50 markers and the BLUP residuals. FDR at different LOD thresholds was calculated by dividing the average number of peaks obtained from five permutations of segregant identities by the number of peaks observed in the real data. We also tested for interactions between each locus with significant additive effects (identified as described in the preceding section) and the rest of the genome in the same manner as for the full two-dimensional scan. We refer to this as the marginal scan. FDR was calculated as above.

Results from the BLUP residual approach were compared with a simpler two-locus interaction model from 'scantwo' in R/QTL (ref. 33) that compares the likelihood ratio of a model that includes an interaction term to a model without this term. From the BLUP residual approach we detected 205 QTL–QTL in the full scan and 266 in the marginal scan. Using the same FDR procedure, 73 QTL–QTL were detected using R/QTL with the full two-dimensional scan and 112 were detected in the marginal scan. All of the R/QTL QTL–QTL interactions were also detected as statistically significant in our BLUP residual models.

Fraction of variance captured by marker subsets. To estimate the fraction of additive variance captured by significant additive QTL, we fit the model $y = \beta X + Za + Zp + e$, where a was calculated from the relatedness of segregants only at the genome-wide significant QTL peak markers for the given trait (A_{QTL}) such that $a \sim N(0, \sigma_{A\_QTL}^2 A_{QTL})$, $p \sim N(0, \sigma_R^2 I_n)$ and $e \sim N(0, \sigma_{EV}^2 I_m)$, and compared it with the same model but with a calculated using the relatedness at all markers in the genome (A) as described above, such that $a \sim N(0, \sigma_A^2 A)$.

We partitioned the interaction variance in a similar manner. Starting with $y = \beta X + Za + Zi + Zp + e$, where $a \sim N(0, \sigma_A^2 A)$ and $i \sim N(0, \sigma_{AA}^2 A \circ A)$, we replaced $A \circ A$ with various subsets of marker combinations. We fit a model with $i \sim N(0, \sigma_{A_{QTL}*A_{QTL}}^2 A_{QTL} \circ A_{QTL})$ to capture the fraction of variance due to all pairwise interactions between significant additive QTL. We fit a model with $i \sim N(0, \sigma_{A_{QTL}*A}^2 A_{QTL} \circ A)$ to capture the fraction of variance due to all pairwise interactions between significant additive QTL and the genome. We fit models

with, $i \sim N\left(0, \sigma_{QQ}^2, QQ\right)$, where σ_{QQ}^2 is the fraction of phenotypic variance captured by significant QTL–QTL interactions and QQ is the relatedness matrix calculated from an $n \times q$ matrix where n is the number of segregants and each column corresponds to the product of the genotypes at the peak markers for genome-wide significant interacting QTL–QTL. The median fraction of interaction variance explained by significant QTL–QTL interactions was calculated as the median of $\sigma_{QQ}^2/\sigma_{AA}^2$ for the given trait.

To estimate the fraction of variance explained by non-specific background linkage effects, N markers or pairs of markers were chosen per trait, where N was the observed number of QTL or QTL–QTL for that trait. GRMs were calculated as above, but for the random marker subsets instead of QTL peak markers. To make this analysis more tractable, phenotype replicates were averaged for each strain and the repeatability term was excluded from the model. Variance components were estimated for each of the models listed above for 50 random draws of N markers for each trait. The median fraction of variance explained from these simulations is plotted in Supplementary Fig. 5.

The individual QTL and QTL–QTL interaction effect sizes shown in Fig. 4 were computed using the analysis of variance function in R with a trait specific multiple regression linear model with all the trait specific significant QTL peak markers and the product of QTL–QTL pair peak markers as fixed effects.

Simulation of genetic architectures. We simulated phenotypes from a range of genetic architectures to test whether the mixed model will appropriately partition variance into additive and interaction components given our experimental design and our observed genotype data. Specifically, we simulated all combinations of either 0, 1, 5, 10, 50 or 500 QTL and/or QTL–QTL interactions. We set the broad-sense heritability (defined for these simulations as additive plus pairwise interaction variance) to 0.75 and varied the additive heritability to range from 0 to 0.75 in increments of 0.15 for all unique combinations of QTL and QTL–QTL interactions. QTL were given equal effects, but the sign of their effect was chosen at random. The positions of additive QTL were chosen randomly for each simulation. The positions of QTL–QTL interactions were chosen from the set of all combinations of additive QTL, but if the target number of QTL–QTL interactions was greater than the set of all combinations of additive QTL, then additional QTL–QTL interaction positions were chosen where neither position had a marginal additive effect. The summed effects of the additive loci were scaled to have the target additive variance and the summed effects of the interacting loci were scaled to have the target interaction variance and these were added to create vector g. Error variance was added from a normal distribution with mean 0 and s.d. $= (1 - H2)/H2 \times \mathrm{var}(g))$. Additive and interacting variance components were estimated with GRMs constructed from all the markers, as described above (Supplementary Data 1). We observed very large estimation errors in case of architectures dominated by one very large-effect interaction, but note that we did not observe such architectures for the traits studied here.

References

1. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4,** e1000008 (2008).
2. Mäki-Tanila, A. & Hill, W. G. Influence of gene interaction on complex trait variation with multilocus models. *Genetics* **198,** 355–367 (2014).
3. Nelson, R. M., Pettersson, M. E. & Carlborg, Ö. A century after Fisher: time for a new paradigm in quantitative genetics. *Trends Genet.* **29,** 669–676 (2013).
4. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc. Natl Acad. Sci. USA* **109,** 1193–1198 (2012).
5. Mackay, T. F. C. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* **15,** 22–33 (2014).
6. Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T.-L. V. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature* **494,** 234–237 (2013).
7. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer, 1998).
8. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157,** 1819–1829 (2001).
9. Rönnegård, L., Pong-Wong, R. & Carlborg, O. Defining the assumptions underlying modeling of epistatic QTL using variance component methods. *J. Hered.* **99,** 421–425 (2008).
10. Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M. & Price, A. L. Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* **46,** 100–106 (2014).
11. Shen, X. The curse of the missing heritability. *Front. Genet.* **4,** 225 (2013).
12. Young, A. I. & Durbin, R. Estimation of epistatic variance components and heritability in founder populations and crosses. *Genetics* **198,** 1405–1416 (2014).
13. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461,** 747–753 (2009).
14. Costanzo, M. *et al.* The genetic landscape of a cell. *Science* **327,** 425–431 (2010).
15. Gerke, J., Lorenz, K. & Cohen, B. Genetic interactions between transcription factors cause natural variation in yeast. *Science* **323,** 498–501 (2009).
16. Gaertner, B. E., Parmenter, M. D., Rockman, M. V, Kruglyak, L. & Phillips, P. C. More than the sum of its parts: a complex epistatic network underlies natural variation in thermal preference behavior in Caenorhabditis elegans. *Genetics* **192,** 1533–1542 (2012).
17. Carlborg, O., Jacobsson, L., Ahgren, P., Siegel, P. & Andersson, L. Epistasis and the release of genetic variation during long-term selection. *Nat. Genet.* **38,** 418–420 (2006).
18. Leamy, L. J., Gordon, R. R. & Pomp, D. Sex-, diet-, and cancer-dependent epistatic effects on complex traits in mice. *Front. Genet.* **2,** 71 (2011).
19. Carlborg, O. *et al.* A global search reveals epistatic interaction between QTL for early growth in the chicken. *Genome Res.* **13,** 413–421 (2003).
20. Carlborg, O., Hocking, P. M., Burt, D. W. & Haley, C. S. Simultaneous mapping of epistatic QTL in chickens reveals clusters of QTL pairs with similar genetic effects on growth. *Genet. Res.* **83,** 197–209 (2004).
21. Carlborg, O., Brockmann, G. A. & Haley, C. S. Simultaneous mapping of epistatic QTL in DU6i x DBA/2 mice. *Mamm. Genome* **16,** 481–494 (2005).
22. Zhu, Z. *et al.* Dominance genetic variation contributes little to the missing heritability for human complex traits. *Am. J. Hum. Genet.* **96,** 377–385 (2015).
23. Bhatia, G. *et al.* Haplotypes of Common SNPs can Explain Missing Heritability of Complex Diseases. doi:10.1101/022418 Preprint at http://biorxiv.org/content/early/2015/07/12/022418, (2015).
24. R Development Core Team, R. R: a language and environment for statistical computing. *R Found. Stat. Comput.* **1,** 409 (2014).
25. Aronesty, E. ea-utilsAvailable at: <http://code.google.com/p/ea-utils> (2011).
26. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
27. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
28. Albert, F. W., Treusch, S., Shockley, A. H., Bloom, J. S. & Kruglyak, L. Genetics of single-cell protein abundance variation in large yeast populations. *Nature* **506,** 494–497 (2014).
29. Pau, G., Fuchs, F., Sklyar, O., Boutros, M. & Huber, W. EBImage--an R package for image processing with applications to cellular phenotypes. *Bioinformatics* **26,** 979–981 (2010).
30. Endelman, J. B. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* **4,** 250–255 (2011).
31. Gilmour, A. R., Thompson, R. & Cullis, B. R. Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51,** 1440–1450 (1995).
32. Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138,** 963–971 (1994).
33. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).

Author contributions

Experiments were designed by J.S.B. and L.K. Experiments were performed by J.S.B. and I.K. The genotyping protocol was developed by S.T. and I.K. Analyses were conducted by J.S.B. The manuscript was written by J.S.B. and L.K. and incorporates comments by M.J.S., F.W.A. and S.T.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Bloom, J. S. *et al.* Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nat. Commun.* 6:8712 doi: 10.1038/ncomms9712 (2015).