I: Derivation of MLEs for Multiplicative and Minimum Epistasis

For multiplicative epistasis, the Maximum Likelihood (ML) estimators for the epistasis coefficient and fitness variance of double mutants are

\[
\hat{\epsilon}^{(p)}_{ij, \text{MLE}} = \frac{w_{ij} w_i w_j}{w_i^2 w_j^2} - 1
\]

\[
\hat{\sigma}^2_{ij, \text{MLE}}^{(p)} = \frac{1}{N} \sum_{n=1}^{N} \left( w_{ij}^n - w_i^n w_j^n (1 + \hat{\epsilon}^{(p)}_{ij, \text{MLE}}) \right)^2
\]

where the right superscript \((p)\) indicates that the estimator is for multiplicative epistasis.

For minimum epistasis, the ML estimators for the epistasis coefficient and fitness variance of double mutants are

\[
\hat{\epsilon}^{(m)}_{ij, \text{MLE}} = \frac{1}{N} \sum_{n=1}^{N} \left( w_{ij}^n - \min(w_i^n, w_j^n) \right)
\]

\[
\hat{\sigma}^2_{ij, \text{MLE}}^{(m)} = \frac{1}{N} \sum_{n=1}^{N} \left( w_{ij}^n - \min(w_i^n, w_j^n) - \hat{\epsilon}^{(m)}_{ij, \text{MLE}} \right)^2
\]

where the right superscript \((m)\) indicates that the estimator is for minimum epistasis. Formulas for the additive MLEs are in the main text.

II: Simulation of Double Mutant Fitness Values

When simulating double mutant fitness values, the support is \([0, 1.2]\), allowing for both synergistic and antagonistic epistasis. If \(g_i\) and \(g_j\) act multiplicatively, that is, there is no multiplicative epistasis, \(w_{ij} \sim N(w_i w_j, \sigma_{ij}^2)_{[0,1.2]}\), and if \(g_i\) and \(g_j\) act according to the minimum definition of
independence, $w_{ij} \sim N(\min(w_i, w_j), \sigma_{ij}^2)[0,1]$. (Distribution for under additive independence is shown in the main text). For multiplicative epistasis, $w_{ij} \sim N(w_i w_j (1 + \epsilon_{ij}), \sigma_{ij}^2)[0,1]$, and for minimum epistasis, $w_{ij} \sim N(\min(w_i, w_j) + \epsilon_{ij}, \sigma_{ij}^2)[0,1]$. (Distribution for additive epistasis is given in the main text).

### III: Details of Analysis of Experimental Data

The first dataset (ST ONGE et al., 2007) included 26 non-essential genes known to confer resistance to the DNA-damaging agent methyl methanesulfonate (MMS), and contained double-deletion strains for all pairs that were possible to construct (a total of 323 double mutant strains). The authors measured the fitnesses of single and double mutants in media both with and without MMS. For both media types, while single mutants have approximately 10 replicate fitness measurements, double mutants have only two replicates.

We modify our procedure to analyze this dataset. For each gene in each pair, we randomly partition its single-mutant fitness values into two parts and use the mean of each part as the two single-mutant fitness replicates. Because of this procedure’s variability, we repeat this procedure 1000 times, each time determining the subtype of epistasis for each pair. We then select gene pairs for which one of the three epistatic subtypes was chosen in at least 900 of the replicates, and select as the final epistatic subtype for each pair that which was chosen in the majority of the replicates. Changing the cutoff value to either 950 or 800 replicates to create this “high-confidence” epistatic set does not greatly affect our results. We average values of the LRT statistic ($\lambda$) and $\epsilon$ over replicates where the given subtype was selected to obtain the final values.

When examining pairs of genes which share “specific” functional links (defined in main text), we obtain the appropriate Gene Ontology (GO) terms in all three GO categories (CONSORTIUM, 2000) for the genes of interest with AmiGO (for ST ONGE et al. (2007) data, GO database release 2009-09-08; for JASNOS and KORONA (2007) data, GO database release 2009-09-22). In order to assess significance of both specific functional links and shared GO Slim terms among the inferred epistatic pairs, we implement both Fisher’s Exact Test and a permutation procedure with 5000 iter-
ations. For each permutation, we randomly partition the $n$ gene pairs originally tested for epistasis into two groups, $x$ pairs and $n - x$ pairs, where $x$ is the number of inferred epistatic pairs. We then count the number out of $x$ pairs that share either a functional link or GO Slim term link. We set the p-value as the proportion of iterations where the number of links exceeds our observation.

To examine only experimentally-verified interactions in the BIOGRID database (STARK et al., 2006), we ignored interactions inferred quantitatively from phenotype data and whose “evidence codes” were one the following: “Dosage Growth Defect,” “Dosage Lethality,” “Dosage Rescue,” “Phenotypic Enhancement,” “Phenotypic Suppression,” “Synthetic Defect,” “Synthetic Haploinsufficiency,” “Synthetic Lethality,” and “Synthetic Rescue.” We focused only on experimentally-validated (physical) interactions, identified by co-localization, co-purification, etc.

In addition, we note that in our formulation (see Section 2.1, main text), the multiplicative epistatic model is undefined when double mutants have a fitness of 0 ($\hat{\epsilon}_{ij} = 0$ and $\hat{\sigma}_{ij}^2 = 0$). In our current implementation of the method, we do not attempt to further subclassify these interactions as either additive or minimum epistasis. We feel that because synthetic lethals have a double mutant fitness of 0, any further subclassification is unnecessary.

**IV: Additional Analyses of Simulated Data**

We assess the accuracy of inferring the correct epistatic subtype when decreasing the sample size to 10 (see Table 2, main text). The ability to distinguish null models from epistatic models is high, with accuracy of approximately 0.85 ~ 0.9 for various values of $\epsilon$. While the accuracy of inferring minimum epistasis decreases slightly (to about 80%), inference of multiplicative epistasis is heavily affected by a reduction of sample size (accuracy of 0.4 ~ 0.6). The inference of additive epistasis is robust to a decrease in replicates when $\epsilon$ is negative; however, when $\epsilon$ is positive, accuracy decreases substantially (0.5 ~ 0.6).

We also relaxed our restriction to deleterious mutations to assess the effect of a more complex fitness landscape on the estimation of epistasis. As illustrated in Table S1, our method is robust to beneficial mutations, which do not decrease the accuracy of inferring the correct epistatic subtype.
If beneficial mutations are of specific interest, in the future it may be worthwhile to consider the Log model as an epistatic subtype (MANI et al., 2008; SANJUAN and ELENA, 2006), as its exclusion may only be warranted for deleterious mutations. We also note that when at least one of the single mutant fitness values in a pair is 1, the predictions under each null model are equivalent and we have no power to detect the correct epistatic subtype (not shown).

Another practical issue is that the accuracy of constructing an epistatic network (indicating all pairwise interactions) decreases substantially as the number of genes increases. Our method is effective in detecting the existence and type of epistasis for one gene pair (accuracy is \( \sim 0.95 \) for minimum epistasis; Table 1, main text). However, for more than 10 genes we must select the epistatic subtype for more than 45 pairs, decreasing accuracy to at most \( 0.95^{45} \approx 0.1 \). This issue arises in all large-scale studies of epistasis, and is not unique to our method. To improve performance, we advocate controlling the FDR.

V: Additional Analyses of Experimental Data

To select one among the three possible multiple-hypothesis testing procedures for the ST ONGE et al. (2007) dataset, we examine the number of inferred epistatic pairs which share “specific” functional links (described in main text). 36 of the 323 pairs examined in the dataset share a specific functional link (including five synthetic lethals in the dataset). When using the Bonferroni correction, out of the 25 gene pairs for which we identify epistasis, only 1 gene pair has a functional link. (See Discussion in main text for implications of this result.) Implementing the FDR procedure (BENJAMINI and HOCHBERG, 1995), we find 188 gene pairs with epistasis (not including synthetic lethal double mutants), of which 20 have a functional link. ST ONGE et al. (2007) discover 133 gene pairs with significant epistasis; as they find this to be a reasonable number, 188 pairs also appears reasonable. Finally, with the pFDR procedure (STOREY, 2002), we find 222 gene pairs with epistasis; this procedure rejects neutrality for all pairs for which an epistatic model was originally selected (for the pFDR, the value above which p-values are estimated to be uniformly distributed is 0.1). In order to assess significance of these values, we perform Fisher’s exact test
and a permutation procedure (described in main text and in Section III); however, the number of functional links observed among epistatic pairs is not significant for any of the multiple-testing measures.

We show several examples of epistatic networks found with our method which considers all three epistatic subtypes and their corresponding null models (with the FDR procedure) when examining the ST ONGE et al. (2007) dataset. In Figure S6(b), we show the epistatic interactions inferred between all pairs of genes with functional links, as well as the type of epistasis for each connection. (For a complete list of all determined epistatic pairs, see Supplemental Results.) We find a mixture of positive and negative epistasis, and slightly different results than the original authors. In Figure S6(c), we show, following the original authors, the homologous recombination pathway including the genes RAD51, RAD52, RAD54, RAD55, and RAD57. While ST ONGE et al. (2007) find 10 epistatic pairs among these genes, we find only 7 epistatic pairs, all of which are classified as minimum epistasis. In Figure S6(d), we present epistasis found for genes of the Shu complex (SHU1, SHU2, CSM2, and PSY3), as do ST ONGE et al. (2007). We do not find PSY3 to interact with any of the genes in this pathway, and find a mixture of multiplicative and minimum epistasis. In constructing the “high-confidence” set of epistatic pairs through sample replications (Section 2.3, main text), we are likely more conservative than ST ONGE et al. (2007) in identifying epistasis. In addition, these authors do not implement any multiple-hypothesis testing procedures. This probably slightly increases the power of their method over ours, yet unfortunately at the expense of detecting more false positives of gene pairs with epistasis.

We also searched for over-representation of GO Slim terms in genes involved in each additive, multiplicative, and minimum subtypes, as well as positive and negative epistasis; however, we found no significant results (CONSORTIUM (2000), www.yeastgenome.org).

We next examine the JASNOS and KORONA (2007) dataset and look for “specific” functional links that exist among epistatic pairs inferred with each of the three multiple hypothesis testing procedures. We examine Gene Ontology (GO) (CONSORTIUM, 2000) terms with fewer than 200 genes associated with them (described in main text) and find 25 links among all 636 tested pairs.
Using the Bonferroni procedure, 56 epistatic pairs are significant, of which 5 have a functional link. The FDR procedure results in 352 significant epistatic pairs, of which 19 have a functional link. For the pFDR, we detect 471 epistatic pairs, of which 21 have a functional link. When assessing significance of the number of functional links among epistatic pairs by permutation (see Section 2.3 (main text) and Section III), we find that only the FDR results in a significant number of functional links (p-value = 0.024).

The epistatic networks discovered with our method and the FDR procedure in the JASNOS and KORONA (2007) dataset are small and isolated, likely due to the examination of only a subset of the possible pairwise genotypes in the study. We annotate functions of the genes in the networks according to the yeast Gene Ontology (GO) Slim Biological Process terms (following JASNOS and KORONA (2007), see Section 2.3 (main text)). Four examples of constructed epistatic networks denoting pairwise epistatic interactions are shown in Figure S7. Figure S7(b) shows minimum epistatic interactions between three genes involved in the biological processes of mitochondrion organization and translation, all of which are functionally linked (as defined above) with the GO term mitochondrial protein formation and synthesis (GO:0032543); MRPL37 and MSK1 are also linked through the mitochondria (GO:005739). These functional links suggest that the epistatic interactions we detect may in fact be meaningful.

In Figure S7(c), we show that genes IFM1 and SNT309 have alleviating positive epistasis; we also find a positive minimum epistatic interaction between SNT309 and BUD28. We find a larger linked group of six genes, among which there are two pairs of genes that have epistasis and share a functional link (Figure S7(d)). In this network, we find a mixture of both minimum and additive epistasis, and both positive and negative epistasis. The functional links and shared GO Slim annotations in this network suggest that this novel epistasis is likely biologically significant.

Finally, Figure S7(e) illustrates an example of a larger group of twelve genes connected by three types of epistasis. Of the twelve genes, four (RDH54, HCM1, CIN8, and SGS1) have roles in the cell cycle. It is possible that other genes in this network, such as SWA2, may also be involved in this process, although they have not yet been studied. A specific functional link also exists between
genes SGS1 and CIN8, which are both involved in mitotic chromosome segregation. We believe that these, along with the other interactions identified (provided in Supplemental Results), are true epistatic interactions.

Although we search for significant enrichment of particular GO terms amongst genes involved in additive, multiplicative, and minimum epistasis, we fail to find any, indicating that the subtypes of epistasis are likely distributed over many functional classes of genes. This lack of significance is also likely due to the small number of double mutants examined.

As discussed in the main text (Discussion), we assess the normality of the replicate fitness data for both studied datasets. A deviation from normality may suggest that the $\chi^2$ approximation for the distribution of the likelihood ratio test statistic under the null hypothesis is inappropriate in some cases. Double-mutant fitness values of the ST ONGE et al. (2007) dataset certainly deviate from normality, as only two replicates are measured (Figure S8). In Figure S8 we also show Q-Q plots for several of the genes of the Shu complex, CSM2, CSM3, and PSY3 (presented in ST ONGE et al. (2007)). Although functional links exist, we do not find PSY3 to interact with either gene; this could be due to the deviation of the CSM3 single-mutant fitness values from normality. For the JASNOS and KORONA (2007) dataset, we examine the Q-Q plots of the genes NCS2 and SGO1 (Figure S9); although a previous interaction has been identified for this pair, we do not detect epistasis. The deviation from normality of the single-mutant fitness values of SGO1 and the small number of double-mutant fitness replicates could explain our inability to detect this interaction. The distributions of fitness values for several other genes show similar deviations from normality (not shown).

VI: Consideration of Alternative Model Selection Procedures

To explore whether the accuracy of inferring the epistatic subtype for pairs of genes could be improved, we also tried using two alternatives to the BIC procedure. The first method is similar to cross-validation. The sample was split randomly into a modeling set and a testing set. We fit each of the six models to the modeling data set and computed the mean squared error per model
for the testing set. After repeating this procedure 1000 times, we selected the model with the smallest accumulated mean squared error. Overall, this method failed to outperform the BIC (data not shown).

The difficulty with testing for epistasis is that the null hypothesis can be any one of the three null models of independence, and is thus not well-specified. We explored using the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) for this problem in order perform a different likelihood ratio test. Under the null model of no epistasis, double mutants can be distributed as a mixture of the three null models; under the epistatic model, the distribution is a mixture model of the three epistatic subtypes. Using the EM algorithm, we maximized the likelihood under each model, and performed a likelihood ratio test. If the null hypothesis of independence was rejected, we selected the epistatic measure for the gene pair as the mixture component that had the highest estimated proportion. However, we found that this method resulted in poorer performance (data not shown). Thus, it appears that the more simplistic BIC method currently implemented is the most appropriate method for this estimation problem.

VII: Supplemental Results

We provide three supplemental files, which are the complete lists of significant epistatic pairs for the three datasets we analyze (using the BIC procedure which considers all three epistatic subtypes and their corresponding null models). The first is the ST ONGE et al. (2007) dataset, measured in the presence of MMS, the second is the ST ONGE et al. (2007) dataset, measured in the absence of MMS, and the third is the JASNOS and KORONA (2007) dataset. Column 1 is an index (from the initial datasets), columns 2 and 3 contain the genes of the epistatic pair, column 4 indicates the epistatic subtype (defined below), column 5 is the likelihood ratio test statistic ($\lambda$), column 6 is the p-value, and column 7 is the estimated value of epistasis ($\epsilon$). Epistatic subtypes are coded as 2 (additive epistasis), 3 (multiplicative epistasis), and 4 (minimum epistasis).
References


