DOI: 10.1534/genetics.109.111120

On the Classification of Epistatic Interactions

Hong Gao,*[,1] Julie M. Granka[†,1] and Marcus W. Feldman[†,2]

**Department of Genetics and [†]Department of Biology, Stanford University School of Medicine, Stanford, California 94305*

Manuscript received October 20, 2009
Accepted for publication December 20, 2009

ABSTRACT

Modern genomewide association studies are characterized by the problem of "missing heritability." Epistasis, or genetic interaction, has been suggested as a possible explanation for the relatively small contribution of single significant associations to the fraction of variance explained. Of particular concern to investigators of genetic interactions is how to best represent and define epistasis. Previous studies have found that the use of different quantitative definitions for genetic interaction can lead to different conclusions when constructing genetic interaction networks and when addressing evolutionary questions. We suggest that instead, multiple representations of epistasis, or epistatic "subtypes," may be valid within a given system. Selecting among these epistatic subtypes may provide additional insight into the biological and functional relationships among pairs of genes. In this study, we propose maximum-likelihood and model selection methods in a hypothesis-testing framework to choose epistatic subtypes that best represent functional relationships for pairs of genes on the basis of fitness data from both single and double mutants in haploid systems. We gauge the performance of our method with extensive simulations under various interaction scenarios. Our approach performs reasonably well in detecting the most likely epistatic subtype for pairs of genes, as well as in reducing bias when estimating the epistatic parameter (ε). We apply our approach to two available data sets from yeast (*Saccharomyces cerevisiae*) and demonstrate through overlap of our identified epistatic pairs with experimentally verified interactions and functional links that our results are likely of biological significance in understanding interaction mechanisms. We anticipate that our method will improve detection of epistatic interactions and will help to unravel the mysteries of complex biological systems.

UNDERSTANDING the nature of genetic interactions is crucial to obtaining a more complete picture of complex biological systems and their evolution. The discovery of genetic interactions has been the goal of many researchers studying a number of model systems, including but not limited to *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Escherichia coli* (You and Yin 2002; Burch *et al.* 2003; Burch and Chao 2004; Tong *et al.* 2004; Drees *et al.* 2005; Sanjuán *et al.* 2005; Segre *et al.* 2005; Pan *et al.* 2006; Zhong and Sternberg 2006; Jasnos and Korona 2007; St. Onge *et al.* 2007; Decourty *et al.* 2008). Recently, high-throughput experimental approaches, such as epistatic mini-array profiles (E-MAPs) and genetic interaction analysis technology for *E. coli* (GIANT-coli), have enabled the study of epistasis on a large scale (Schuldiner *et al.* 2005, 2006; Collins *et al.* 2006, 2007; Typas *et al.* 2008). However, it remains unclear whether the computational and statistical methods currently in use to identify these interactions are indeed the most appropriate.

The study of genetic interaction, or "epistasis," has had a long and somewhat convoluted history. Bateson (1909) first used the term epistasis to describe the ability of a gene at one locus to "mask" the mutational influence of a gene at another locus (Cordell 2002). The term "epistacy" was later coined by Fisher (1918) to denote the statistical deviation of multilocus genotype values from an additive linear model for the value of a phenotype (Phillips 1998, 2008).

These origins are the basis for the two main current interpretations of epistasis. The first, as introduced by Bateson (1909), is the "biological," "physiological," or "compositional" form of epistasis, concerned with the influence of an individual's genetic background on an allele's effect on phenotype (Cheverud and Routman 1995; Phillips 1998, 2008; Cordell 2002; Moore and Williams 2005). The second interpretation, attributed to Fisher, is "statistical" epistasis, which in its linear regression framework places the phenomenon of epistasis in the context of a population (Wagner *et al.* 1998; Wade *et al.* 2001; Wilke and Adami 2001; Moore and Williams 2005; Phillips 2008). Each of these approaches is equally valid in studying genetic interactions; however, confusion still exists about how to best reconcile the methods and results of the two (Phillips 1998, 2008; Cordell 2002; Moore and Williams 2005; Liberman and Feldman 2006; Aylor and Zeng 2008).

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.109.111120/DC1.

[1]These authors contributed equally to this work.

[2]*Corresponding author:* 478 Herrin Hall, Stanford University, Stanford, CA 94305. E-mail: mfeldman@stanford.edu

Aside from the distinction between the statistical and the physiological definitions of epistasis, inconsistencies exist when studying solely physiological epistasis. For categorical traits, physiological epistasis is clear as a "masking" effect. When noncategorical or numerical traits are measured, epistasis is defined as the deviation of the phenotype of the multiple mutant from that expected under independence of the underlying genes.

The "expectation" of the phenotype under independence, that is, in the absence of epistasis, is not defined consistently between studies. For clarity, consider epistasis between pairs of genes and, without loss of generality, consider fitness as the phenotype. The first commonly used definition of independence, originating from additivity, defines the effect of two independent mutations to be equal to the sum of the individual mutational effects. A second, motivated by the use of fitness as a phenotype, defines the effect of the two mutations as the product of the individual effects (Elena and Lenski 1997; Desai *et al.* 2007; Phillips 2008). A third definition of independence has been referred to as "minimum," where alleles at two loci are independent if the double mutant has the same fitness as the less-fit single mutant. Mani *et al.* (2008) claim that this has been used when identifying pairwise epistasis by searching for synthetic lethal double mutants (Tong *et al.* 2001, 2004; Pan *et al.* 2004, 2006; Davierwala *et al.* 2005). A fourth is the "Log" definition presented by Mani *et al.* (2008) and Sanjuan and Elena (2006). The less-frequently used "scaled ε" (Segre *et al.* 2005) measure of epistasis takes the multiplicative definition of independence with a scaling factor.

These different definitions of independence are partly due to distinct measurement "scales." For some traits, a multiplicative definition of independence may be necessary to identify epistasis between two genes, whereas for other traits, additivity may be appropriate (Falconer and Mackay 1995; Wade *et al.* 2001; Mani *et al.* 2008; Phillips 2008). An interaction found under one independence definition may not necessarily be found under another, leading to different biological conclusions (Mani *et al.* 2008).

Mani *et al.* (2008) suggest that there may be an "ideal" definition of independence for all gene pairs for identifying functional relationships. However, it is plausible that different representations of independence for two genes may reflect different biological properties of the relationship (Kupper and Hogan 1978; Rothman *et al.* 1980). "Two categories of general interest [the additive and multiplicative definitions, respectively] are those in which etiologic factors act interchangeably in the same step in a multistep process, or alternatively act at different steps in the process" (Rothman *et al.* 1980, p. 468). In some cases, the discovery of epistasis may merely be an artifact of using an incorrect null model (Kupper and Hogan 1978). It may be necessary to represent "independence" differently, resulting in different statistical measures of interactions, for different pairs of genes depending on their functions.

Previous studies have suggested that different pairs of loci may have different modes of interaction and have attempted to subclassify genetic interactions into regulatory hierarchies and mutually exclusive "interaction subtypes" to elucidate underlying biological properties (Avery and Wasserman 1992; Drees *et al.* 2005; St. Onge *et al.* 2007). We suggest that epistatic relationships can be divided into several subtypes, or forms, corresponding to the aforementioned definitions of independence. As a particular gene pair may deviate from independence according to several criteria, we do not claim that these subtypes are necessarily mutually exclusive. We attempt to select the most likely epistatic subtype that is the best statistical representation of the relationship between two genes. To further subclassify interactions, epistasis among deleterious mutations can take one of two commonly used forms: positive (equivalently alleviating, antagonistic, or buffering) epistasis, where the phenotype of the double mutant is less severe than expected under independence, and negative (equivalently aggravating, synergistic, or synthetic), where the phenotype is more severe than expected (Segre *et al.* 2005; Collins *et al.* 2006; Desai *et al.* 2007; Mani *et al.* 2008).

Another objective of such distinctions is to reduce the bias of the estimator of the epistatic parameter (ε), which measures the extent and direction of epistasis for a given gene pair. Mani *et al.* (2008), assuming that the overall distribution of ε should be centered around 0, find that inaccurately choosing a definition of independence can result in increased bias when estimating ε. For example, using the minimum definition results in the most severe bias when single mutants have moderate fitness effects, and the additive definition results in the largest positive bias when at least one gene has an extreme fitness defect (Mani *et al.* 2008). Therefore, it is important to select an optimal estimator for ε for each pair of genes from among the subtypes of epistatic interactions.

Epistasis may be important to consider in genomic association studies, as a gene with a weak main effect may be identified only through its interaction with another gene or other genes (Frankel and Schork 1996; Culverhouse *et al.* 2002; Moore 2003; Cordell 2009; Moore and Williams 2009). Epistasis has also been studied extensively in the context of the evolution of sex and recombination. The mutational deterministic hypothesis proposes that the evolution of sex and recombination would be favored by negative epistatic interactions (Feldman *et al.* 1980; Kondrashov 1994); many other studies have also studied the importance of the form of epistasis (Elena and Lenski 1997; Otto and Feldman 1997; Burch and Chao 2004; Keightley and Otto 2006; Desai *et al.* 2007; MacCarthy and Bergman 2007). Indeed, according to Mani *et al.* (2008,

p. 3466), "the choice of definition [of epistasis] alters conclusions relevant to the adaptive value of sex and recombination."

Given fitness data from single and double mutants in haploid organisms, we implement a likelihood method to determine the subtype that is the best statistical representation of the epistatic interaction for pairs of genes. We use maximum-likelihood estimation and the Bayesian information criteria (BIC) (Schwarz 1978) with a likelihood-ratio test to select the most appropriate null or epistatic model for each putative interaction. We conduct extensive simulations to gauge the performance of our method and demonstrate that it performs reasonably well under various interaction scenarios. We apply our method to two data sets with fitness measurements obtained from yeast (Jasnos and Korona 2007; St. Onge *et al.* 2007), whose authors assume only multiplicative epistasis for all interactions. By examining functional links and experimentally validated interactions among epistatic pairs, we demonstrate that our results are biologically meaningful. Studying a random selection of genes, we find that minimum epistasis is more prevalent than both additive and multiplicative epistasis and that the overall distribution of ε is not significantly different from zero (as Jasnos and Korona 2007 suggest). For genes in a particular pathway, we advise selecting among fewer epistatic subtypes. We believe that our method of epistatic subtype classification will aid in understanding genetic interactions and their properties.

METHODS

We consider epistasis among L genes (represented by $g_1, g_2, \ldots, g_L$) given the phenotypes of each of the possible single- and double-mutant haplotypes, each with a given number of replicates. The phenotype we study is fitness scaled by wild type, obtained either through experiments or through simulation (Schuldiner *et al.* 2005, 2006; Segre *et al.* 2005; Collins *et al.* 2006, 2007; Jasnos and Korona 2007; St. Onge *et al.* 2007). The fitness of the wild-type genotype g_0 is 1, and the relative fitnesses of single mutants g_i ($1 \le i \le L$) and double mutants g_{ij} ($1 \le i < j \le L$) are denoted by w_i and w_{ij}, respectively. Supposing that we have N replicates of fitness values for each mutant, the superscript n represents measurements in the nth replicate ($1 \le n \le N$). We denote by σ_i^2 and σ_{ij}^2 the variances in fitnesses of single mutants g_i and double mutants g_{ij}, respectively. We define ε_{ij} as the epistatic coefficient of a genetic interaction between genes g_i and g_j.

Epistatic subtype selection: Pairwise physiological epistasis for measured numerical traits is defined as the deviation of the observed phenotype of the double mutant from the phenotype expected when the alleles at each locus act independently. As mentioned above, the additive and multiplicative models are widely used to represent independence, while the minimum model is used by investigators searching for synthetic lethal double mutants (Mani *et al.* 2008). Thus, we focus predominantly on these three models, as do Mani *et al.* (2008). The Log model is practically equivalent to the multiplicative model for deleterious mutations (Mani *et al.* 2008); since we generally focus on deleterious mutations, we do not consider this model. We do not include a model based on the scaled ε measure of epistasis (Segre *et al.* 2005), as it was found to be less meaningful than the multiplicative model in describing interactions (St. Onge *et al.* 2007).

We thus use six models: the three models of independence (additive, multiplicative, and minimum) and the three models (subtypes) of epistasis (additive, multiplicative, and minimum epistasis), which incorporate a deviation from the independence models. We have three major questions. First, when data follow one of the three models of independence, do we incorrectly infer the presence of epistasis? Second, can we detect epistasis between two genes if they interact? Third, can we accurately determine which subtype is the most appropriate for a particular pair of genes if they interact?

The three null models are as follows. If genes g_i and g_j act additively, *i.e.*, there is no additive epistasis, the null model for the double-mutant fitness in the nth experiment (w_{ij}^n) is a normal distribution with mean $w_i^n + w_j^n - 1$ and variance σ_{ij}^2. If g_i and g_j act multiplicatively, w_{ij}^n is a normal random variable with mean $w_i^n w_j^n$ and variance σ_{ij}^2. The third model specifying the absence of epistasis assumes that w_{ij}^n is normal with mean $\min(w_i^n, w_j^n)$ and variance σ_{ij}^2. We emphasize that the minimum model described here is a model of independence.

In the presence of additive epistasis ε_{ij}, the probability distribution of the double-mutant fitness w_{ij} in the nth experiment, given w_i^n, w_j^n, and variance σ_{ij}^2, is

$$\mathbb{P}(w_{ij}^n = x \mid w_i^n, w_j^n, \varepsilon_{ij}, \sigma_{ij}^2) \sim N(w_i^n + w_j^n - 1 + \varepsilon_{ij}, \sigma_{ij}^2). \quad (1)$$

When g_i and g_j exhibit multiplicative epistasis ε_{ij}, w_{ij}^n has probability distribution

$$\mathbb{P}(w_{ij}^n = x \mid w_i^n, w_j^n, \varepsilon_{ij}, \sigma_{ij}^2) \sim N(w_i^n w_j^n (1 + \varepsilon_{ij}), \sigma_{ij}^2). \quad (2)$$

This form of multiplicative epistasis is slightly different from that of other authors, who suggest $w_{ij} = w_i w_j + \varepsilon_{ij}$ (Segre *et al.* 2005; Jasnos and Korona 2007; St. Onge *et al.* 2007; Mani *et al.* 2008); our formulation is simply a rescaling of this formula. Our motivation for Equation 2 is that it relates more closely to the large body of research representing the fitness of an individual with k mutations as $w_k = e^{-sk^{1+\varepsilon}}$ (Elena and Lenski 1997; Otto and Feldman 1997; Wilke and Adami 2001; Desai *et al.* 2007).

In the case of minimum epistasis, w_{ij}^n is also distributed normally:

$$\mathbb{P}(w_{ij}^n = x \mid w_i^n,\, w_j^n,\, \varepsilon_{ij},\, \sigma_{ij}^2) \sim N(\min(w_i^n,\, w_j^n) + \varepsilon_{ij}, \sigma_{ij}^2). \quad (3)$$

An epistatic subtype refers to one of the three epistatic models shown in the equations above.

To calculate the likelihood of the observed double-mutant fitnesses, we assume that each pair of genes is independent of all other pairs; *i.e.*, there are no interactions of higher order. The log likelihood of the data based on N replicates is

$$\begin{aligned}\log L = \sum_{n=1}^{N} \sum_{(i,j):1\le i<j\le L} & \log \mathbb{P}_{\mathrm{E}}(w_{ij}^n = x \mid w_i^n, w_j^n, \varepsilon_{ij}, \sigma_{ij}^2) \times \\ & \times I_{\{g_i \text{ interacts with } g_j\}} \\ & + \log \mathbb{P}_{\mathrm{N}}(w_{ij}^n = x \mid w_i^n,\, w_j^n,\, \sigma_{ij}^2) \times (1 - I_{\{g_i \text{ interacts with } g_j\}}),\end{aligned} \quad (4)$$

where $I_{\{\Phi\}}$ is an indicator function and is equal to one if Φ is true and zero otherwise, and $\mathbb{P}_{\mathrm{E}}$ and $\mathbb{P}_{\mathrm{N}}$ are the probabilities (given above) under the epistatic and null models, respectively.

Since we assume that each pair of genes is independent of all other pairs, we study each separately. We use the BIC (Schwarz 1978), a commonly used model selection criterion, to determine the most likely interaction subtype, if any, of each pair. To calculate the BIC, we first estimate the parameters ε_{ij} and σ_{ij}^2 using maximum-likelihood estimation (MLE). For additive epistasis, the ML estimators (MLEs) for ε_{ij} and σ_{ij}^2 are

$$\begin{aligned}\hat{\varepsilon}_{ij_{\mathrm{MLE}}}^{(a)} &= \overline{w_{ij}} - \overline{w_i} - \overline{w_j} + 1 \\ \hat{\sigma}_{ij_{\mathrm{MLE}}}^{2(a)} &= \frac{1}{N}\sum_{n=1}^{N}(w_{ij}^n - w_i^n - w_j^n + 1 - \hat{\varepsilon}_{ij_{\mathrm{MLE}}}^{(a)})^2,\end{aligned} \quad (5)$$

where $\overline{w_{\cdot}}$ denotes the mean fitness of genotype $g_{\cdot}$ over all N replicates, and the right superscript (a) indicates that the estimator is for additive epistasis. The MLEs for multiplicative and minimum epistasis are given in supporting information, File S1.

We select the model that is the best statistical representation of the interaction for each pair of genes using the BIC, defined as

$$\begin{aligned}\mathrm{BIC} &= n_{\mathrm{p}}\log(N) - 2\log L_{\mathrm{M}} \\ &= n_{\mathrm{p}}\log(N) + N(1 + \log(2\pi\hat{\sigma}_{ij_{\mathrm{MLE}}}^2)),\end{aligned} \quad (6)$$

where L_{M} and n_{p} are the maximum likelihood and the number of free parameters, respectively, in the given model. For the three null models, $n_{\mathrm{p}} = 1$, as only σ_{ij}^2 is estimated; for the three epistatic models, $n_{\mathrm{p}} = 2$, as both σ_{ij}^2 and ε_{ij} are estimated. The model with the smallest BIC is selected as optimal, rewarding models with a higher likelihood but penalizing for additional parameters.

For each pair, we perform a likelihood-ratio test (LRT) of the epistatic model with the lowest BIC against its corresponding null model as an optional step to assess the significance of epistasis. The LRT statistic is given as

$$\lambda = -2[\log L_{\mathrm{M_{null}}} - \log L_{\mathrm{M_{epistatic}}}]. \quad (7)$$

Under the null distribution, λ is approximately χ^2-distributed with 1 d.f. (for the one additional parameter ε estimated in the epistatic model).

When identifying epistasis for many gene pairs, the problem of simultaneously testing multiple hypotheses arises. We consider three commonly used multiple-testing corrections, which use the *P*-values obtained from the LRT above. The first is the Bonferroni correction, which sets the chosen significance level divided by the number of tests (in our case, the number of putatively epistatic gene pairs) as a *P*-value cutoff. This is generally considered to be the most conservative procedure to control the familywise error rate (the probability of rejecting at least one true null hypothesis among all tests). The second multiple-testing correction, proposed by Benjamini and Hochberg (1995), controls the false discovery rate (FDR, or the expected proportion of null hypotheses that are falsely rejected). Third, we control the positive false discovery rate (pFDR), which is slightly more lenient than the FDR procedure, as suggested by Storey (2002). We advocate determining significance using the FDR, which is intermediate between the other two procedures (see application to experimental data: *Estimation of epistatic measures* and application to experimental data). We implement our methods in R (www.r-project.org, code available upon request).

Generation of simulated data: We carried out extensive simulations to assess the performance of our method. For simplicity, we assume that each mutation is deleterious; *i.e.*, the fitnesses of single mutants cannot be greater than that of wild type ($w_0 = 1$). (In Table S1, we consider relaxing this assumption to include beneficial mutations). We sample the fitnesses of single mutants g_i and g_j from a truncated normal distribution on [0, 1] with mean μ and standard deviation σ; *i.e.*, $w_i \sim N(\mu_i, \sigma_i^2)_{[0,1]}$ and $w_j \sim N(\mu_j, \sigma_j^2)_{[0,1]}$. Although many deletion strains found in nature have a wild-type fitness, we simulate with this support to encompass the entire range of fitness values possible for deleterious mutants. We choose the support for double-mutant fitnesses to be [0, 1.2]; since both synergistic and antagonistic epistasis may be possible, the fitness of double mutants may be >1. We simulate w_{ij} given w_i and w_j following the definitions presented in methods: *Epistatic subtype selection.* For example, if genes g_i and g_j act additively and independently, $w_{ij} \sim N(w_i + w_j - 1, \sigma_{ij}^2)_{[0,1.2]}$; under additive epistasis, $w_{ij} \sim N(w_i + w_j - 1 + \varepsilon_{ij}, \sigma_{ij}^2)_{[0,1.2]}$. Distributions under multiplicative and minimum null and epistatic models are given in File S1.

Analysis of experimental data: We analyze two data sets of fitness measurements in yeast (*S. cerevisiae*) (Jasnos and Korona 2007; St. Onge *et al.* 2007). In addition to choosing the appropriate epistatic subtype on the basis of the BIC, we perform a LRT to test

epistatic models against their corresponding null models and correct for multiple testing with the three methods described in *Epistatic subtype selection*. For all tests, we select an α significance level of 0.05.

The first data set (St. Onge *et al.* 2007) included 26 nonessential 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 (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 ~10 replicate fitness measurements, double mutants have only 2. Since our method assumes the same number of replicates for single and double mutants, we perform our method with two sampled single-mutant fitness measurements. We repeat this sampling 1000 times and select an epistatic subtype for the pairs that were chosen as epistatic in at least 900 of the 1000 replicates (see File S1 for details).

In the second study, Jasnos and Korona (2007) performed 639 random crosses among 758 randomly selected gene deletions to create a series of double mutants and assayed the growth fitnesses of each single and double mutant. As an equal number of replicates are available for both single and double mutants ($n \approx 5$), we apply our method to this data set directly.

We perform several analyses to validate the use of our method. First, we examine the number of epistatic pairs sharing "specific" functional links, a term used by both St. Onge *et al.* (2007) and Tong *et al.* (2004). St. Onge *et al.* (2007) define these to be genes that share Gene Ontology (GO) terms (Gene Ontology Consortium 2000) that have associated with them <30 genes; Tong *et al.* (2004) define the cutoff for specificity to be ≤200 associated genes. We examine our data using these and several intermediate cutoffs. Second, we count epistatic pairs sharing the broader GO Slim terms (obtained from the Yeast Genome Database, www.yeastgenome.org, accessed September 11, 2009). To assess the significance of the number of specific functional links or shared GO Slim terms among inferred epistatic pairs, we implement both Fisher's exact test and a permutation procedure (File S1), which produce comparable results. Third, we examine overlap of inferred epistatic pairs with experimentally verified physical interactions in yeast [BIOGRID database (Stark *et al.* 2006), www.thebiogrid.org, Release 2.0.56; see File S1].

APPLICATION TO SIMULATED DATA

Estimation of epistatic measures: We perform 10,000 simulations, following the methods described in *Generation of simulated data*, under each of the six models (three null models and three epistatic models) for each set of parameters shown in Table 1. For each simulation, we sample 50 fitness replicates ($N = 50$), obtain the BIC for each of the six models, and select the model with the lowest BIC. We compute accuracy as the proportion of simulations in which the model under which the data are simulated is selected.

TABLE 1

Fractions of simulations that recover the true model (null or epistatic) among the six models (three null models and three epistatic models) with different values of the epistasis coefficient (ε_{ij}), fitness standard deviation ($\sigma = \sigma_i = \sigma_j = \sigma_{ij}$), and mean fitnesses of single mutants (μ_i and μ_j)

σ	μ_i	μ_j	$P_N^{(a)}$	$P_N^{(p)}$	$P_N^{(m)}$	ε_{ij}	$P_E^{(a)}$	$P_E^{(p)}$	$P_E^{(m)}$
0.05	0.8	0.8				−0.3	0.996	0.906	0.929
0.05	0.8	0.8				−0.2	0.992	0.902	0.958
0.05	0.8	0.8	0.986	0.983	0.994	−0.1	0.965	0.873	0.981
0.05	0.8	0.8				0.1	0.739	0.720	0.997
0.05	0.8	0.8				0.2	0.591	0.605	0.996
0.05	0.8	0.8				0.3	0.746	0.610	0.999
0.05	0.9	0.6				−0.3	0.999	0.958	0.978
0.05	0.9	0.6				−0.2	0.995	0.939	0.965
0.05	0.9	0.6	0.988	0.987	0.992	−0.1	0.986	0.901	0.972
0.05	0.9	0.6				0.1	0.864	0.832	0.994
0.05	0.9	0.6				0.2	0.830	0.825	0.999
0.05	0.9	0.6				0.3	0.888	0.821	0.999
0.1	0.8	0.8				−0.3	0.978	0.870	0.915
0.1	0.8	0.8				−0.2	0.961	0.866	0.884
0.1	0.8	0.8	0.955	0.952	0.993	−0.1	0.953	0.758	0.896
0.1	0.8	0.8				0.1	0.754	0.711	0.995
0.1	0.8	0.8				0.2	0.689	0.618	0.997
0.1	0.8	0.8				0.3	0.908	0.517	1.00
0.1	0.9	0.6				−0.3	0.780	0.942	0.946
0.1	0.9	0.6				−0.2	0.947	0.910	0.945
0.1	0.9	0.6	0.973	0.971	0.991	−0.1	0.963	0.744	0.915
0.1	0.9	0.6				0.1	0.833	0.731	0.987
0.1	0.9	0.6				0.2	0.804	0.758	0.994
0.1	0.9	0.6				0.3	0.882	0.781	0.997

$P_N^{(a)}$ is the fraction of simulations (simulated under the additive null model) with the additive null model attaining the minimum value of the BIC among the six models. $P_N^{(p)}$ and $P_N^{(m)}$ are defined in a manner similar to $P_N^{(a)}$ for the multiplicative and minimum models, respectively. $P_E^{(a)}$ represents the power of our method, *i.e.*, the percentage of simulations (simulated under the additive epistatic model with epistasis coefficient ε_{ij}) with the additive epistatic model attaining the minimum value of the BIC among the six models. $P_E^{(p)}$ and $P_E^{(m)}$ are defined in a manner similar to $P_E^{(a)}$ for the multiplicative and minimum models, respectively. Fifty replicate fitness values are simulated. See Table S6 for a similar table with 10 replicate values.

Table 1 demonstrates that our method can select the correct null or epistatic model for various parameter combinations. When simulating under one of the three null models, our method identifies the correct model with almost 100% accuracy. Overall, the minimum null model is identified with higher accuracy than either the additive or the multiplicative null models. Identification of the minimum epistatic subtype is consistently the most accurate (near 100%), while the accuracy of selecting the additive and multiplicative epistatic subtypes varies with the epistatic coefficient ε_{ij}.

If two single mutants have the same mean fitness, our method is less able to detect multiplicative epistasis than

either additive or minimum epistasis (Table 1). Introducing a fairly large difference between single-mutant fitness values slightly increases the accuracy of selecting multiplicative epistasis when data are simulated under this definition. For all epistatic models, the probability of selecting the appropriate model generally increases with the absolute value of ε_{ij}: the larger the absolute value of ε_{ij}, the larger is the difference between the expected double-mutant fitness under the correct model and the other models. We also have little ability to consistently distinguish between the additive and multiplicative models. See also Figure S1 and Figure S2.

Testing hypotheses with the LRT and FDR procedures (Benjamini and Hochberg 1995) gives nearly identical accuracy to that described above (results not shown). With data simulated under the null model, as expected, the FDR slightly decreases the false positive rate; for data simulated under an epistastic model, power decreases by at most 1%. Although the false positive rate decreases when using the Bonferroni correction, we obtain slightly lower power than with the FDR; the pFDR (Storey 2002) does not greatly improve power over the FDR (results not shown). The FDR procedure thus seems preferable.

Bias and variance of the epistatic parameter (ε): Mani *et al.* (2008) show that using different definitions of independence results in different estimators of the epistatic parameter (ε), some with larger biases and variances than others. To evaluate the bias and variance of $\hat{\varepsilon}_{\text{MLE}}$ estimated with our method, we perform simulations and compute its mean squared error (MSE) under each epistatic model. Because the FDR procedure does not greatly affect the accuracy of our inference (see above), we do not implement it in these simulations. For the nth simulation, $\text{MSE}_{\varepsilon}^{n} = (\hat{\varepsilon}_{\text{MLE}} - \varepsilon_n)^2$, where ε_n is the true value used in the nth simulation and $\hat{\varepsilon}_{\text{MLE}}$ is estimated under the specific epistatic model. If the simulated $|\varepsilon|$ is large, the MSE of $\hat{\varepsilon}_{\text{MLE}}$ is the smallest and is nearly zero when the true epistatic model is selected (Figure S3). When the simulated $|\varepsilon|$ decreases, we find that while the MSE is always small when the correct model is selected, at times the MSE is even lower under an incorrect model (Figure S3). In general, our method seems to reduce the bias and variance of ε.

APPLICATION TO EXPERIMENTAL DATA

Epistasis has been studied extensively in yeast (*S. cerevisiae*), a model organism for which generating a large number of mutants and examining growth phenotypes are experimentally tractable. Of interest are studies measuring single- and double-mutant fitnesses on a continuous scale. With our proposed method, we reanalyze two data sets by detecting epistasis and its subtype for each pair of mutants (Jasnos and Korona 2007; St. Onge *et al.* 2007).

St. Onge *et al.* (2007) data set: St. Onge *et al.* (2007) examined 26 nonessential genes known to confer resistance to MMS, constructed double-deletion strains for 323 double-mutant strains (all but two of the total possible pairs), and assumed the multiplicative form of epistasis for all interactions (see Methods: *Analysis of experimental data*). Following these authors, we focus on single- and double-mutant fitnesses measured in the presence of MMS. (For results in the absence of MMS, see File S1 and File S1_2.)

Using the resampling method described in *Analysis of experimental data* and File S1, 222 gene pairs pass the cutoff of having epistasis inferred in at least 900 of 1000 replicates. This does not include 5 synthetic lethal gene pairs. Hypothesis testing and a multiple-testing procedure (for 222 simultaneous hypotheses) are necessary to determine the final epistatic pairs.

To select one among the three multiple-testing procedures, we follow St. Onge *et al.* (2007) and examine gene pairs that share specific functional links (see *Analysis of experimental data*). The Bonferroni method is likely too conservative, yielding only 25 significantly epistatic pairs with only one functional link among them; alternatively, the pFDR procedure appears to be too lenient in rejecting independence for all 222 pairs. Therefore, we use the FDR procedure (although the number of functional links is not significant) and detect 193 epistatic pairs, of which 5 (2.6%) are synthetic lethals, 19 (9.8%) have additive epistasis, 33 (17.1%) have multiplicative epistasis, and 136 (70.5%) have minimum epistasis (Table 2, complete list in File S1_1). We find 29 gene pairs with positive (alleviating) epistasis and 159 gene pairs with negative (aggravating) epistasis.

To further validate the use of our method and the FDR procedure, we assess by Fisher's exact test the significance of an enrichment of both Biological Process and all GO Slim term links among epistatic pairs, neither of which are significant (Gene Ontology Consortium 2000; www.yeastgenome.org; Table 3). Our discovered epistatic interactions do overlap with experimentally verified physical interactions [BIOGRID database (Stark *et al.* 2006); Table 3 and Table S4]. Although some of the previously unidentified interactions that we identify could be false positives, many are likely to be new discoveries.

The epistatic subtypes we consider are not necessarily mutually exclusive. To more fully assess the assumptions of our method, we also consider several of the possible subsets of the epistatic subtypes (and their corresponding null models) in our procedure. As the minimum epistatic subtype was the most frequently selected in this data set, we first do not include the minimum null model or the minimum epistatic model in our procedure (*i.e.*, we select from among four rather than six models for a pair; Table 3, column 3). In Table 3, we also examine our method's performance when it selects from only additive epistasis or independence (column 4), only multiplicative epistasis and independence (column 5), or only minimum epistasis and independence (column 6). Allowing for different subsets of

TABLE 2

Summary of gene pairs with the indicated epistatic subtypes, inferred using the FDR procedure with the BIC method that considers all three epistatic subtypes and their corresponding null models

Epistatic subtype	Study S	Study J
All	193 (100%)	352 (100%)
	$\bar{\varepsilon} = -0.060$	$\bar{\varepsilon} = -0.001$
	$\varepsilon_{q_{0.5}} = -0.096$	$\varepsilon_{q_{0.5}} = -0.059$
Additive	19 (9.8%)	35 (9.9%)
	$\bar{\varepsilon} = 0.115$*	$\bar{\varepsilon} = 0.193$***
	$\varepsilon_{q_{0.5}} = 0.131$	$\varepsilon_{q_{0.5}} = 0.188$
Multiplicative	33 (17.1%)	63 (17.9%)
	$\bar{\varepsilon} = 0.048$	$\bar{\varepsilon} = 0.017$
	$\varepsilon_{q_{0.5}} = -0.166$	$\varepsilon_{q_{0.5}} = -0.115$
Minimum	136 (70.5%)	254 (72.2%)
	$\bar{\varepsilon} = -0.111$***	$\bar{\varepsilon} = -0.032$**
	$\varepsilon_{q_{0.5}} = -0.091$	$\varepsilon_{q_{0.5}} = -0.065$

Numbers are the counts of each type, and percentages are given of the total number of epistatic pairs. The mean ($\bar{\varepsilon}$) and median ($\varepsilon_{q_{0.5}}$) of the epistatic parameter (ε) are given for each subtype, with "*" indicating that the mean of ε is significantly different from 0 (*, *P*-value ≤0.05; **, *P*-value ≤0.01; ***, *P*-value ≤0.001). Study S refers to the St. Onge *et al.* (2007) data set, and study J refers to the Jasnos and Korona (2007) data set. (For study S, five of the epistatic pairs are synthetic lethals and are not shown; as a result, percentages do not sum to 100%.)

subtypes does not greatly affect the number of inferred epistatic pairs with experimentally identified interactions (Table 3, Table S4). However, there are a significant number of epistatic pairs with functional links only when the minimum epistatic subtype is not included (also see Table S4 and Table S5). It is not immediately clear which epistatic subtypes are the most appropriate for these data, although including the minimum subtype may not be appropriate (Mani *et al.* 2008) (see discussion).

Although it may be best to consider fewer epistatic subtypes for this specific data set, we report our results including all three epistatic subtypes and their corresponding null models (Table 2). Our method identifies epistasis for 88 of the same pairs as St. Onge *et al.* (2007), although we identify 105 epistatic pairs not identified by the original authors (Figure S4, Table S4). St. Onge *et al.* (2007) find that epistatic pairs with a functional link have a positively shifted distribution of epistasis. We find no such shift in epistasis values (Table 2, Figure S5). We also demonstrate [described in application to simulated data: *Bias and variance of the epistatic parameter* (ε)] that our method seems to reduce bias of the epistatic parameter (ε) (Table 2). [Note that in the absence of MMS, we find that $\bar{\varepsilon}$ is negative and significantly different from 0 (Table S3).] When considering only a subset of the epistatic subtypes, however, we find $\bar{\varepsilon}$ to be positive and significantly different from zero (results not shown). See File S1, Figure S6, and Figure S7 for additional discussion of the epistatic pairs we identify.

Jasnos and Korona (2007) data set: The Jasnos and Korona (2007) data set included 758 yeast gene deletions known to cause growth defects and reports fitnesses of only a sparse subset of all possible gene pairs [≈0.2% of the possible pairwise genotypes, or 639 pairs of $\binom{758}{2}$]. Because the authors do not identify epistatic pairs in a hypothesis-testing framework, we cannot explicitly compare our conclusions with theirs.

To validate our method, we examine gene pairs that have specific functional links (see methods: *Analysis of experimental data*). When defining a functional link using GO terms (Gene Ontology Consortium 2000) with <30 genes associated with them, only 1 of 639 tested gene pairs has a functional link. Raising the threshold of associated genes to 50 and 100, the number of tested pairs with functional links rises only to 3 and 9, respectively. Because of the large number of random genes and the sparse number of gene pairs in this data set, we follow Tong *et al.* (2004) and select GO terms that have associated with them ≤200 genes. Twenty-five of 639 tested pairs then have a functional link.

Only the FDR multiple-testing procedure results in a significant enrichment of functional links among epistatic pairs (Table 3); thus, it is likely the most appropriate of the three multiple-testing methods (full results in File S1). With the FDR procedure we find 352 significant epistatic pairs, of which 35 (9.9%) have additive epistasis, 63 (17.9%) have multiplicative epistasis, and 254 (72.2%) have minimum epistasis (Table 2; for a complete list, see File S1_3). These proportions of inferred subtypes suggest that the authors' original restriction to multiplicative epistasis may be inappropriate. We find 141 gene pairs with positive epistasis and 211 gene pairs with negative epistasis.

We do not find a significant number of epistatic pairs with shared GO Slim Biological Process terms (see *Analysis of experimental data*), but do when considering all shared GO Slim terms (Table 3). However, the significant enrichment of functional links among epistatic pairs suggests that the epistasis found with the BIC and FDR procedures is valid and of biological significance. We also find an overlap of epistatic pairs with experimentally validated physical interactions (from the BIOGRID database).

As with the St. Onge *et al.* (2007) data set, we also consider some of the possible subsets of the three epistatic subtypes (and their corresponding null models) in our model selection procedure (Table 3; see also Table S5). In contrast to the St. Onge *et al.* (2007) data set, using all three epistatic subtypes results in a significant number of epistatic pairs with functional links; this measure is not significant when using any of the other subsets of the subtypes. This suggests that our proposed

TABLE 3

Comparison of validation measures for each data set for different variations of the FDR and BIC procedures, considering only a subset of epistatic subtypes with their corresponding null models: all epistatic subtypes (A, P, and M); only the additive and multiplicative subtypes (A and P); and only the additive (A), only the multiplicative (P), or only the minimum (M) subtype (see text for details)

	Subtypes considered in BIC procedure				
	A, P, M	A, P	A	P	M
Study J					
No. found (636)	352	273	263	231	329
Functional links (25)	19 (0.0255)*	13 (0.2320)	11 (0.4689)	10 (0.4227)	15 (0.2619)
GO Slim terms (Biological Process) (115)	69 (0.1573)	50 (0.4874)	55 (0.0736)	44 (0.3534)	68 (0.04902)*
GO Slim terms (all) (369)	224 (0.0009)*	172 (0.01654)*	160 (0.1297)	146 (0.0273)*	213 (0.0003)*
Experimentally identified (3)	3	2	1	2	3
Study S					
No. found (323)	193	192	247	171	243
Functional links (36)	21 (0.6450)	29 (0.0041)*	34 (0.0031)*	29 (0.0003)*	24 (0.9256)
GO Slim terms (Biological Process) (283)	174 (0.0657)	174 (0.03656)*	223 (0.0010)*	153 (0.1825)	213 (0.5534)
GO Slim terms (all) (307)	185 (0.2866)	182 (0.6926)	237 (0.1472)	162 (0.6997)	231 (0.5908)
Experimentally identified (29)	17	22	24	23	21

Numbers in parentheses indicate *P*-values by Fisher's exact test. "*" indicates significance. Study J refers to the Jasnos and Korona (2007) data set, and study S refers to the St. Onge *et al.* (2007) data set measured in the presence of MMS. Numbers in parentheses indicate the total number of tested pairs and the total number of each type of link found in each complete data set.

method with three epistatic subtypes may indeed be the most appropriate for data sets with randomly selected genes.

We examined the distribution of the estimated values of the epistatic parameter (ε) for all pairs with significant epistasis. Jasnos and Korona (2007), in assuming only multiplicative epistasis, conclude that epistasis is predominantly positive. However, we find that the estimated mean of epistasis is not significantly different from zero (two-sided *t*-test, *P*-value = 0.9578; Figure 1 and Table 2). This also supports that our method reduces bias of the estimator of ε. We report in further detail on the epistatic pairs identified in File S1.

DISCUSSION

We present a likelihood estimation method using the BIC, which, given fitness data from single and double mutants, determines the most appropriate epistatic subtype for pairs of genes that interact. Through extensive simulations, we demonstrate that selection of the epistatic subtype is extremely important and that our method can indeed do so accurately. We demonstrate through the analysis of two data sets that the epistasis we discover is likely biologically meaningful. Besides the detection of genetic interactions and clarification of their functions, accurate assessment of epistasis is also important for evolutionary theory. For example, the distribution of epistatic effects has important implications for the role of epistasis in the evolution of sex and recombination. We demonstrate that our method reduces bias of the epistatic parameter ε.

Our approach is novel for several reasons. While in reality different subtypes and representations of epistasis may be best for different pairs of loci, classical studies have usually assumed only one subtype of epistasis for all pairs. Our method selects the most likely subtype for pairs of genes, providing a more precise picture of the loci actually exhibiting epistasis and insight into their modes of genetic interaction. Second, some past studies of epistasis have not clearly distinguished between the epistatic and independence models. Jasnos and Korona (2007), for example, calculate values of ε without formally assessing whether the perceived epistasis is merely statistical noise. With the presented BIC and LRT procedures, we distinguish between the epistatic and null models in a hypothesis-testing framework and address the multiple-hypothesis testing problem with the FDR (Benjamini and Hochberg 1995).

Applying our method to two yeast data sets (Jasnos and Korona 2007; St. Onge *et al.* 2007), we detect both novel and experimentally verified epistatic interactions. We believe that an enrichment of specific functional links among epistatic pairs is likely the most direct indication of biological importance of inferred epistasis (Tong *et al.* 2004; St. Onge *et al.* 2007), as genetic interactions probably occur on a fine biological scale (see methods: *Analysis of experimental data*).

Including all three epistatic subtypes and their corresponding null models in our procedure results in a significant enrichment of functional links among identified epistatic pairs for the Jasnos and Korona (2007) data set, but not the St. Onge *et al.* (2007) data set (both with and without MMS). Performance for the St. Onge *et al.* (2007) data set is improved when only the

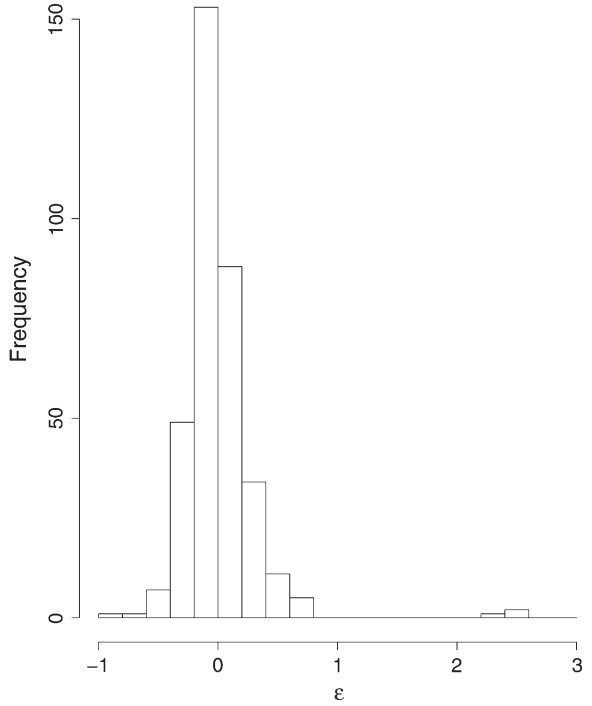

Figure 1.—Distribution of the epistasis values (ε) for significant epistatic pairs in the Jasnos and Korona (2007) data set, determined using the FDR procedure and the BIC method including all three epistatic subtypes and their corresponding null models. Mean of ε is −0.0009, with a standard deviation of 0.3177; median value is −0.0587. A similar plot is shown in Figure 3 of Jasnos and Korona (2007).

additive or the multiplicative subtypes, or both, are considered (Table 3). This apparent difference between data sets may be due to the fact that the St. Onge *et al.* (2007) data set includes genes known to be involved in conferring resistance to MMS, whereas the Jasnos and Korona (2007) data set measured a random selection of genes and gene pairs. In examining epistasis among genes with similar functions, it thus may be more appropriate to consider fewer epistatic subtypes. This difference between the data sets may also be due to the low sample size for double mutants in the St. Onge *et al.* (2007) data set.

We find predominantly minimum epistasis, *i.e.*, a deviation from the minimum null model, when considering all three subtypes in both data sets. This is likely valid for the Jasnos and Korona (2007) data set, but it remains unclear which subtypes are the most appropriate for the St. Onge *et al.* (2007) data set. Mani *et al.* (2008) suggest that the minimum definition of independence may not always be useful. Examining another data set with randomly selected genes and gene pairs will help to assess the performance of our method.

The slight overlap of epistatic pairs inferred using different subsets of the three subtypes emphasizes that the subtypes we select are not necessarily mutually exclusive (Table 3). We claim to select the statistically best-fitting epistatic subtype for pairs of genes to gain additional insight into their mode of interaction. For example, a double mutant can deviate from both additive and multiplicative independence. Selecting the additive epistatic subtype with our method suggests that additive epistasis may be the more appropriate and useful way to represent the relationship between the two genes. The different results obtained considering different subtypes again emphasizes the importance of selecting among them (Mani *et al.* 2008).

While Jasnos and Korona (2007) find a skew toward positive epistasis, we find estimates of epistasis to be centered around zero (Table 2). For both data sets, however, the mean of additive epistasis (ε) is positive and significantly different from 0, whereas the mean of minimum epistasis (ε) is negative and significantly different from 0 (Table 2). Applying our method to data sets with a larger number of genes may provide a more accurate view of the types and strengths of epistasis often found in nature.

Inference about epistatic subtypes is affected by many factors, including the number of replicate experiments and the variance of mutant fitnesses. Accuracy clearly increases with the number of replicate measurements, but in practice <10 replications are performed (Jasnos and Korona 2007; St. Onge *et al.* 2007). We assess our method's accuracy when the number of replicates is decreased from 50 to 10 and to 2 (Table S2 and Table S6). As expected, accuracy decreases, but varies depending on the type of epistasis examined. Our failure to identify epistasis for several experimentally identified interacting pairs in the St. Onge *et al.* (2007) data set is likely due to our low power with only two double-mutant fitness measurements.

A large variance in mutant fitnesses also decreases the accuracy of our method. Our simulations use 0.05 as the standard deviation of mutant fitnesses, which is the typical value found for single and double mutants in the Jasnos and Korona (2007) data set. To assess the effect of variance of fitness on accuracy, we ran simulations with the standard deviation increased to 0.1; this slightly reduces the ability to distinguish the true epistatic model from the others (Table 1).

We also address the approximation of the distribution of the LRT statistic (λ) to the χ^2_1-distribution under the null hypothesis. For both data sets, very few functional links exist among epistatic pairs inferred with the Bonferroni method. This indicates that epistatic pairs with functional links have somewhat larger *P*-values and are significant only under less stringent multiple-testing protocols. The χ^2-approximation may not always be appropriate, as the measured fitnesses sometimes deviate from normality in both data sets (Figure S8 and Figure S9). This may also explain our inability to detect some experimentally validated interactions. We urge measuring a large number of replicates to meet the normality requirement and to increase the power of identifying epistatic interactions.

Our method is not limited to studying fitness as a phenotype and can be applied to any trait, although it is necessary to scale the phenotype of interest by the wild-type value. Our likelihood-based method is applicable only to studies of experimental data where multiple replicates are measured; it is not immediately applicable to studies of computationally predicted fitness values, with only one measurement (Segre *et al.* 2005). However, adding random noise terms to each predicted value to create multiple "replicates" could allow our method to be applied in the future. In addition, it is important to note that the epistasis identified with this method is a property of the alleles at each locus; it is conceivable that when examining other pairs of alleles at the tested loci, different results could be obtained. We also note that our method is applicable only to haploid systems. Identifying epistasis and constructing double mutants is much more difficult in diploid systems, as more complex statistical issues arise due to dominance and other effects (Otto and Feldman 1997; Wagner *et al.* 1998; Gregersen *et al.* 2006; Li *et al.* 2007; Flint and Mackay 2009). Since current methods to identify epistasis in haploid systems are still in their infancy, it is important to first develop and perfect these methods before addressing diploid systems.

Our method is one in a wide research field studying genetic interactions and in a long history of the study of epistasis (Hartman and Tippery 2004; Bansal *et al.* 2007; Roguev *et al.* 2008; Chipman and Singh 2009). As mentioned above, our method leaves many open questions. Not all of the genetic interactions that we find may translate into physical interactions, and many researchers have pursued the question of how genetic and protein–protein interactions are related (Lee *et al.* 2004; Wong *et al.* 2005; Schwartz *et al.* 2009). The translation of discoveries of epistasis into discoveries of functional interactions is a very intriguing, albeit complex, problem. We expect that our method will shed light on intrinsic properties of epistasis and may help to address these broader issues.

We thank Elhanan Borenstein for his helpful comments and suggestions, as well as two anonymous reviewers and the editor of Genetics. This research project is supported by National Institutes of Health (NIH) grant GM28016 to M.W.F., NIH grant R01 GM073059 to Hua Tang, and a Stanford Dean's postdoctoral fellowship to H.G. J.M.G. is supported by the Stanford Genome Training Program, which is funded by the NIH (academic year 2008–2009), and by the National Science Foundation Graduate Research Fellowship (2009 to present).

LITERATURE CITED

Avery, L., and S. Wasserman, 1992 Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet. **8:** 312–316.

Aylor, D., and Z. Zeng, 2008 From classical genetics to quantitative genetics to systems biology: modeling epistasis. PLoS Genet. **4:** e1000029.

Bansal, M., V. Belcastro, A. Ambesi-Impiombato and D. di Bernardo, 2007 How to infer gene networks from expression profiles. Mol. Syst. Biol. **3:** 78.

Bateson, W., 1909 *Mendel's Principles of Heredity.* Cambridge University Press, Cambridge, UK/London/New York.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B **57:** 289–300.

Burch, C., and L. Chao, 2004 Epistasis and its relationship to canalization in the RNA virus *phi6.* Genetics **167:** 559–567.

Burch, C., P. Turner and K. Hanley, 2003 Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design. J. Evol. Biol. **16:** 1223–1235.

Cheverud, J. M., and E. J. Routman, 1995 Epistasis and its contribution to genetic variance components. Genetics **139:** 1455–1461.

Chipman, K. C., and A. J. Singh, 2009 Predicting genetic interactions with random walks on biological networks. BMC Bioinformatics **10:** 17.

Collins, S. R., M. Schuldiner, N. J. Krogan and J. S. Weissman, 2006 A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. Genome Biol. **7:** R63.

Collins, S. R., K. M. Miller, N. L. Maas, A. Roguev, J. Fillingham *et al.*, 2007 Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature **446:** 806–810.

Cordell, H. J., 2002 Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum. Mol. Genet. **11:** 2463–2468.

Cordell, H. J., 2009 Detecting gene-gene interactions that underlie human diseases. Nat. Rev. Genet. **10:** 392–404.

Culverhouse, R., B. Suarez, J. Lin and T. Reich, 2002 A perspective on epistasis: limits of models displaying no main effect. Am. J. Hum. Genet. **70:** 461–471.

Davierwala, A., J. Haynes, Z. Li, R. Brost, M. D. Robinson *et al.*, 2005 The synthetic genetic interaction spectrum of essential genes. Nat. Genet. **37:** 1147–1152.

Decourty, L., C. Saveanu, K. Zemam, F. Hantraye, E. Frachon *et al.*, 2008 Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. Proc. Natl. Acad. Sci. USA **105:** 5821–5826.

Desai, M., D. Weissman and M. Feldman, 2007 Evolution can favor antagonistic epistasis. Genetics **177:** 1001–1010.

Drees, B. L., V. Thorsson, G. W. Carter, A. W. Rives, M. Z. Raymond *et al.*, 2005 Derivation of genetic interaction networks from quantitative phenotype data. Genome Biol. **6:** R38.

Elena, S., and R. Lenski, 1997 Test of synergistic interactions among deleterious mutations in bacteria. Nature **390:** 395–398.

Falconer, D., and T. Mackay, 1995 *Introduction to Quantitative Genetics.* Longman, Harlow, UK.

Feldman, M., F. Christiansen and L. Brooks, 1980 Evolution of recombination in a constant environment. Proc. Natl. Acad. Sci. USA **77:** 4838–4841.

Fisher, R., 1918 The correlations between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinb. **52:** 399–433.

Flint, J., and T. F. C. Mackay, 2009 Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. **19:** 723–733.

Frankel, W., and N. Schork, 1996 Who's afraid of epistasis? Nat. Genet. **14:** 371–373.

Gene Ontology Consortium, 2000 Gene ontology: tool for the unification of biology. Nat. Genet. **25:** 25–29.

Gregersen, J. W., K. R. Kranc, X. Ke, P. Svendsen, L. S. Madsen, *et al.*, 2006 Functional epistasis on a common MHC haplotype associated with multiple sclerosis. Nature **443:** 574–577.

Hartman, J., and N. Tippery, 2004 Systematic quantification of gene interactions by phenotypic array analysis. Genome Biol. **5:** R49.

Jasnos, L., and R. Korona, 2007 Epistatic buffering of fitness loss in yeast double deletion strains. Nat. Genet. **39:** 550–554.

Keightley, P. D., and S. P. Otto, 2006 Interference among deleterious mutations favours sex and recombination in finite populations. Nature **443:** 89–92.

Kondrashov, A., 1994 Muller's ratchet under epistatic selection. Genetics **136:** 1469–1473.

Kupper, L., and M. Hogan, 1978 Interaction in epidemiologic studies. Am. J. Epidemiol. **108:** 447–453.

Lee, I., S. V. Date, A. T. Adai and E. M. Marcotte, 2004 A probabilistic functional network of yeast genes. Science **306:** 1555–1558.

Li, H., G. Gao, J. Li, G. P. Page and K. Zhang, 2007 Detecting epistatic interactions contributing to human gene expression using the CEPH family data. BMC Proc. **1**(Suppl. 1):S67.

Liberman, U., and M. W. Feldman, 2006 Evolutionary theory for modifiers of epistasis using a general symmetric model. Proc. Natl. Acad. Sci. USA **103:** 19402–19406.

MacCarthy, T., and A. Bergman, 2007 Coevolution of robustness, epistasis, and recombination favors asexual reproduction. Proc. Natl. Acad. Sci. USA **104:** 12801–12806.

Mani, R., R. P. St. Onge, J. L. I. Hartman, G. Giaever and F. P. Roth, 2008 Defining genetic interaction. Proc. Natl. Acad. Sci. USA **105:** 3461–3466.

Moore, J., 2003 The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum. Hered. **56:** 73–82.

Moore, J., and S. Williams, 2005 Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. BioEssays **27:** 637–646.

Moore, J., and S. Williams, 2009 Epistasis and its implications for personal genetics. Am. J. Hum. Genet. **85:** 309–320.

Otto, S. P., and M. Feldman, 1997 Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. **51:** 134–147.

Pan, X., D. Yuan, D. Xiang, X. Wang, S. Sookhai-Mahadeo *et al.*, 2004 A robust toolkit for functional profiling of the yeast genome. Mol. Cell **16:** 487–496.

Pan, X., P. Ye, D. Yuan, X. Wang, J. Bader *et al.*, 2006 A DNA integrity network in the yeast S*accharomyces cerevisiae*. Cell **124:** 1069–1081.

Phillips, P., 1998 The language of gene inetraction. Genetics **149:** 1167–1171.

Phillips, P., 2008 Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. **9:** 855–867.

Roguev, A., S. Bandyopadhyay, M. Zofall, K. Zhang, T. Fischer *et al.*, 2008 Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science **322:** 405–410.

Rothman, K., S. Greenland and A. Walker, 1980 Concepts of interaction. Am. J. Epidemiol. **112:** 467–470.

Sanjuán, R., and S. Elena, 2006 Epistasis correlates to genomic complexity. Proc. Natl. Acad. Sci. USA **103:** 104402–104405.

Sanjuán, R., J. Cuevas, A. Moya and S. Elena, 2005 Epistasis and the adaptability of an RNA virus. Genetics **170:** 1001–1008.

Schuldiner, M., S. Collins, N. Thompson, V. Denic, A. Bhamidipati *et al.*, 2005 Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell **123:** 507–519.

Schuldiner, M., S. Collins, J. Weissman and N. Krogan, 2006 Quantitative genetic analysis in S*accharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. Methods **40:** 344–352.

Schwartz, A. S., J. Yu, K. R. Gardenour, R. L. Finley, Jr. and T. Ideker, 2009 Cost-effective strategies for completing the interactome. Nat. Methods **6:** 55–61.

Schwarz, G., 1978 Estimating the dimension of a model. Ann. Stat. **6:** 461–464.

Segre, D., A. DeLuna, G. Church and R. Kishony, 2005 Modular epistasis in yeast metabolism. Nat. Genet. **37:** 77–83.

St. Onge, R. P., R. Mani, J. Oh, M. Proctor, E. Fung *et al.*, 2007 Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. Nat. Genet. **39:** 199–206.

Stark, C., B. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz *et al.*, 2006 Biogrid: a general repository for interaction datasets. Nucleic Acids Res. **34:** D535–D539.

Storey, J., 2002 A direct approach to false discovery rates. J. R. Stat. Soc. B **64**(part 3): 479–498.

Tong, A., M. Evangelista, A. Parsons, H. Xu, G. Bader *et al.*, 2001 Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science **294:** 2364–2368.

Tong, A., G. Lesage, G. Bader, H. Ding, H. Xu *et al.*, 2004 Global mapping of the yeast genetic interaction network. Science **303:** 808–813.

Typas, A., R. Nichols, D. Siegele, M. Shales, S. R. Collins *et al.*, 2008 High-throughput quantitative analyses of genetic interactions in E. coli. Nat. Methods **5:** 781–787.

Wade, M., R. Winther, A. Agrawal and C. Goodnight, 2001 Alternative definitions of epistasis: independence and interaction. Trends Ecol. Evol. **16:** 498–504.

Wagner, G. P., M. D. Laubichler and H. Bagheri-Chaichian, 1998 Genetic measurement theory of epistatic effects. Genetics **102:** 569–580.

Wilke, C. O., and C. Adami, 2001 Interaction between directional epistasis and average mutational effects. Proc. R. Soc. Lond. Ser. B Biol. Sci. **268:** 1469–1474.

Wong, S., L. Zhang and F. P. F. Roth, 2005 Discovering functional relationships: biochemistry versus genetics. Trends Genet. **21:** 424–427.

You, L., and J. Yin, 2002 Dependence of epistasis on environment and mutation severity as revealed by *in silico* mutagenesis of phage T7. Genetics **160:** 1273–1281.

Zhong, W., and P. Sternberg, 2006 Genome-wide prediction of C. *elegans* genetic interactions. Science **311:** 1481–1484.

Communicating editor: P. C. Phillips

GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.109.111120/DC1

On the Classification of Epistatic Interactions

Hong Gao, Julie M. Granka and Marcus W. Feldman

DOI: 10.1534/genetics.109.111120

FILE S1
SUPPORTING INFORMATION

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

$$\begin{aligned} \hat{\epsilon}^{(p)}_{ij_{MLE}} &= \frac{\overline{w_{ij}w_iw_j}}{\overline{w_i^2w_j^2}} - 1 \\ \hat{\sigma^2}^{(p)}_{ij_{MLE}} &= \frac{1}{N}\sum_{n=1}^{N}(w_{ij}^n - w_i^n w_j^n(1+\hat{\epsilon}^{(p)}_{ij_{MLE}}))^2, \end{aligned} \tag{S1}$$

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

$$\begin{aligned} \hat{\epsilon}^{(m)}_{ij_{MLE}} &= \frac{1}{N}\sum_{n=1}^{N}(w_{ij}^n - \min(w_i^n, w_j^n)) \\ \hat{\sigma^2}^{(m)}_{ij_{MLE}} &= \frac{1}{N}\sum_{n=1}^{N}(w_{ij}^n - \min(w_i^n, w_j^n) - \hat{\epsilon}^{(m)}_{ij_{MLE}})^2, \end{aligned} \tag{S2}$$

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, $w_{ij} \sim N(w_iw_j, \sigma^2_{ij})_{[0,1.2]}$, and if g_i and g_j act according to the minimum definition of

independence, $w_{ij} \sim N(\min(w_i, w_j), \sigma_{ij}^2)_{[0,1.2]}$. (Distribution for under additive independence is shown in the main text). For multiplicative epistasis, $w_{ij} \sim N(w_i w_j(1+\epsilon_{ij}), \sigma_{ij}^2)_{[0,1.2]}$, and for minimum epistasis, $w_{ij} \sim N(\min(w_i, w_j) + \epsilon_{ij}, \sigma_{ij}^2)_{[0,1.2]}$ (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 (λ) and ϵ 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 ($\hat{\epsilon_{ij}}$ = -1 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 $0.85 \sim 0.9$ for various values of ϵ. 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 $0.4 \sim 0.6$). The inference of additive epistasis is robust to a decrease in replicates when ϵ is negative; however, when ϵ is positive, accuracy decreases substantially ($0.5 \sim 0.6$).

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 ~ 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 χ^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 (λ), column 6 is the p-value, and column 7 is the estimated value of epistasis (ϵ). Epistatic subtypes are coded as 2 (additive epistasis), 3 (multiplicative epistasis), and 4 (minimum epistasis).

References

BENJAMINI, Y. and Y. HOCHBERG, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B **57**: 289–300.

CONSORTIUM, T. G. O., 2000 Gene ontology: tool for the unification of biology. Nat. Genet. **25**: 25–29.

DEMPSTER, A. P., N. M. LAIRD and D. B. RUBIN, 1977 Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society series B **39**: 1–38.

JASNOS, L. and R. KORONA, 2007 Epistatic buffering of fitness loss in yeast double deletion strains. Nat. Genet. **39**: 550–554.

MANI, R., R. P. ST ONGE, J. L. I. HARTMAN, G. GIAEVER and F. P. ROTH, 2008 Defining genetic interaction. Proc. Natl. Acad. Sci. **105**: 3461–66.

SANJUAN, R. and S. ELENA, 2006 Epistasis correlates to genomic complexity. Proc. Natl. Acad. Sci. **103**: 104402–104405.

ST ONGE, R. P., R. MANI, J. OH, M. PROCTOR, E. FUNG *et al.*, 2007 Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. Nat. Genet. **39**: 199–206.

STARK, C., B. BREITKREUTZ, T. REGULY, L. BOUCHER, A. BREITKREUTZ *et al.*, 2006 Biogrid: A general repository for interaction datasets. Nucleic Acids Res. **34**: D535–D539.

STOREY, J., 2002 A direct approach to false discovery rates. J. R. Statist. Soc. B **64, Part 3**: 479–98.

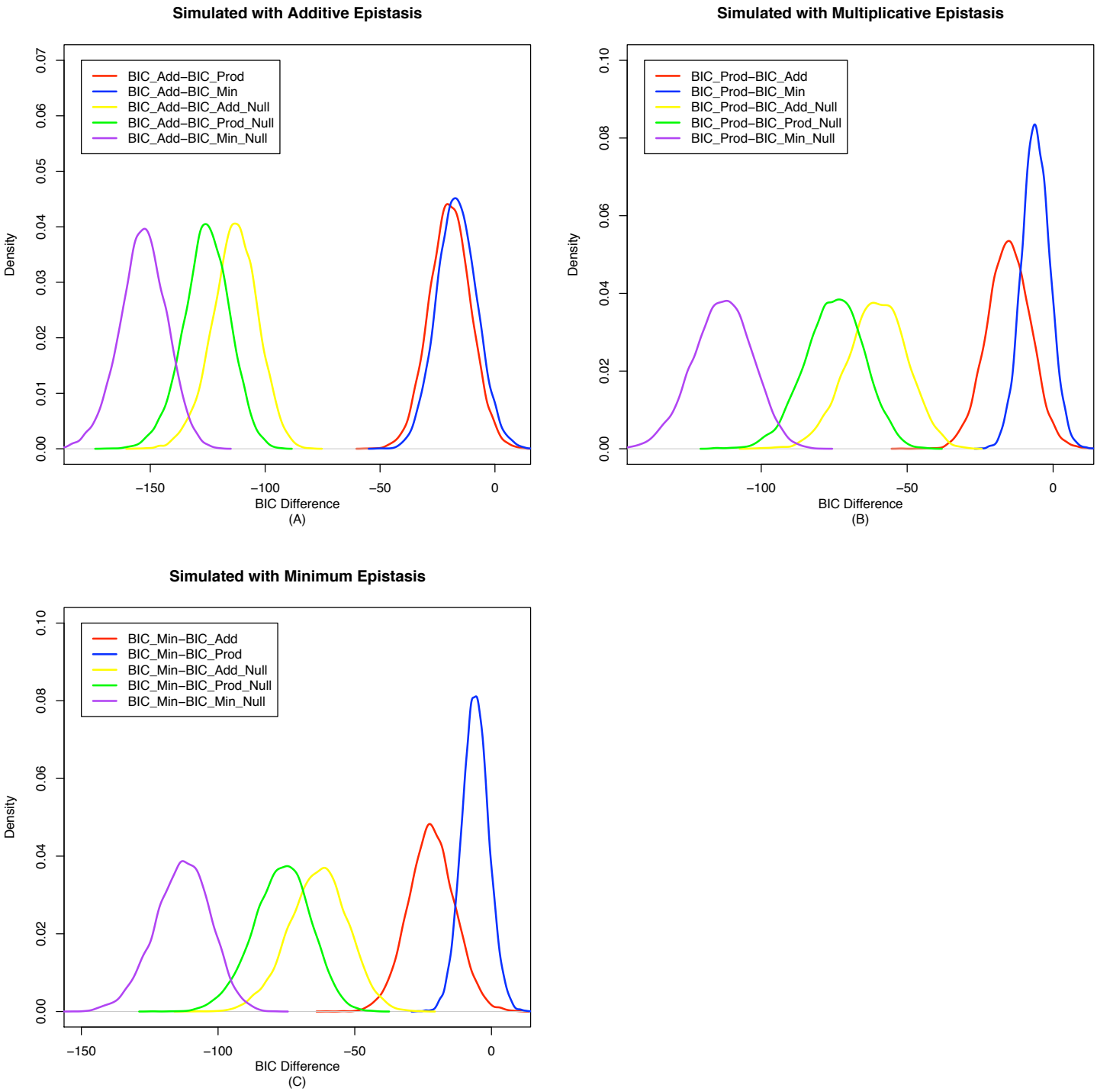

FIGURE S1.—Power of our method to uncover the true (simulated) epistatic model assuming $\varepsilon_{ij} = -0.3$, $\sigma_{ij} = 0.1$, and $\mu_i = \mu_j = 0.8$. (A) The area under each curve and less than zero represents the percentage of simulations with the additive epistatic model favored over the other models shown in the legend. The red, blue, yellow, green, and purple lines represent the comparison of the BIC between the additive epistatic model and the multiplicative epistatic model, minimum epistatic model, additive null model, multiplicative null model, and minimum null model, respectively. (B) and (C) are shown in a manner similar to (A) for the multiplicative and minimum epistatic models, respectively.

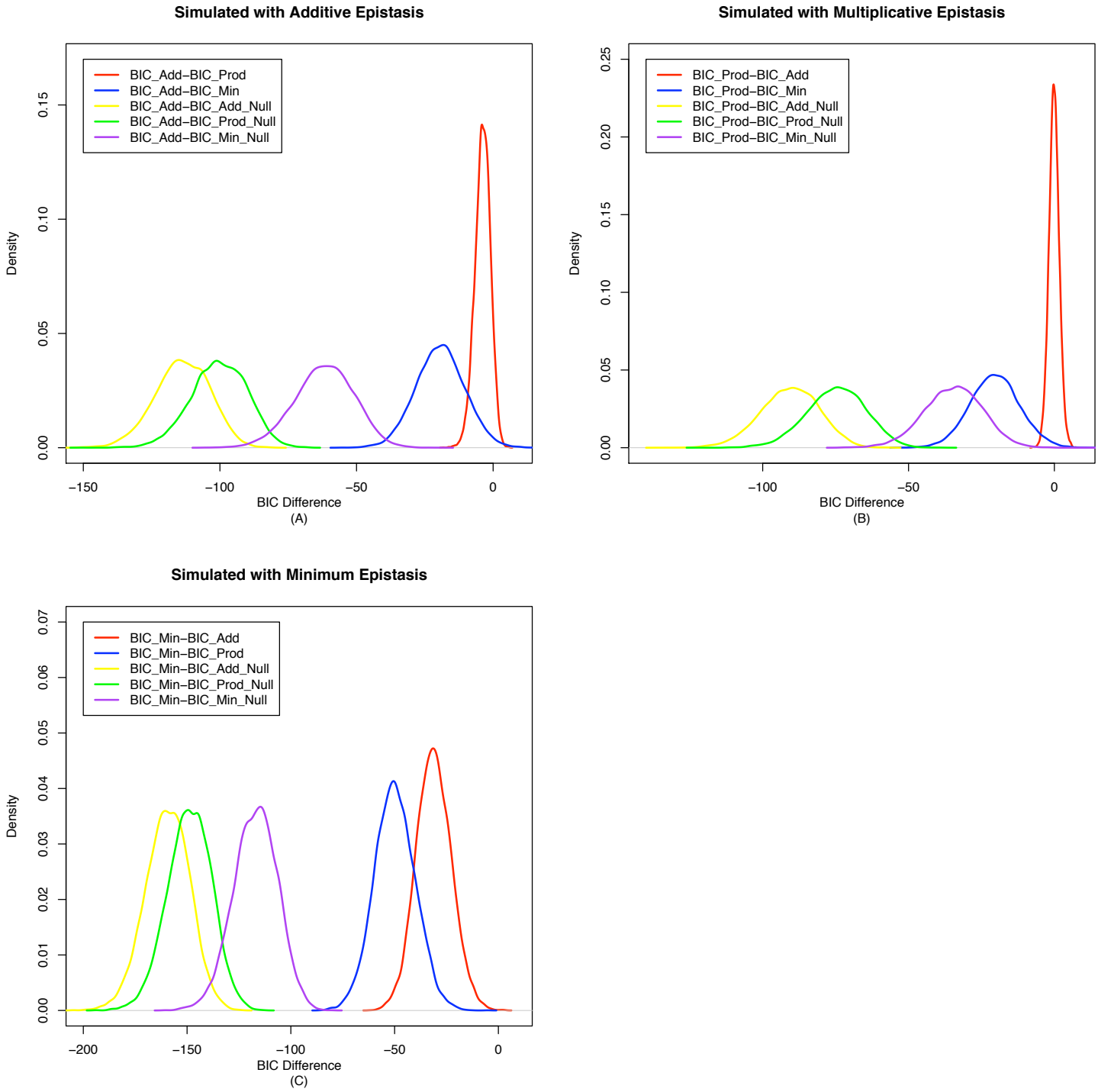

FIGURE S2.—Power of our method to uncover the true (simulated) epistatic model assuming ε_{ij}=0.3, σ_{ij}=0.1, and $\mu_i = \mu_j$ = 0.8. Legends are as defined in Figure S1.

e

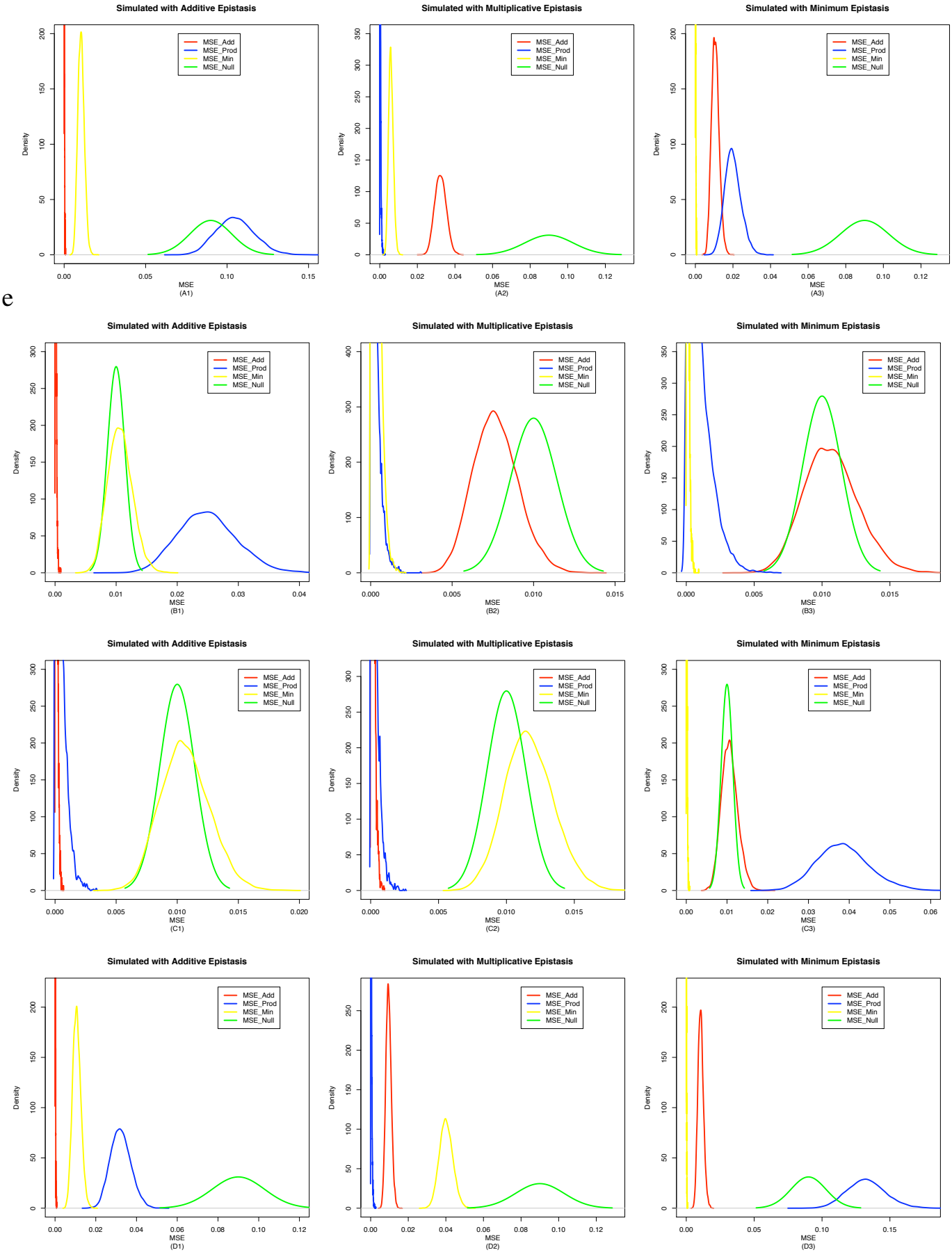
Simulated with Additive Epistasis
Simulated with Multiplicative Epistasis
Simulated with Minimum Epistasis
MSE_Add
MSE_Prod
MSE_Min
MSE_Null
Density
MSE
(A1)
(A2)
(A3)
(B1)
(B2)
(B3)
(C1)
(C2)
(C3)
(D1)
(D2)
(D3)

FIGURE S3.—The distribution of the mean squared error (MSE) of the MLE of the epistatic parameter under the additive, multiplicative, and minimum epistatic models as well as the corresponding null model (with mean fitnesses of single mutants μ_i = 0.9 and μ_j =0.6). Parts A1-3 show the distribution of MSE assuming epistatic coefficient ε_{ij} = -0.3; B1-3 assume ε_{ij} = -0.1; C1-3 assume ε_{ij} = 0.1; D1-3 assume ε_{ij} = 0.3. The red line represents the distribution of the MSE $\hat{\epsilon}_{MLE}$ under the additive epistatic model; the blue line for $\hat{\epsilon}_{MLE}$ under the multiplicative epistatic model; the yellow line for $\hat{\epsilon}_{MLE}$under the minimum epistatic model; the green line for $\hat{\epsilon}_{MLE}$ under the null model (of the simulated definition).

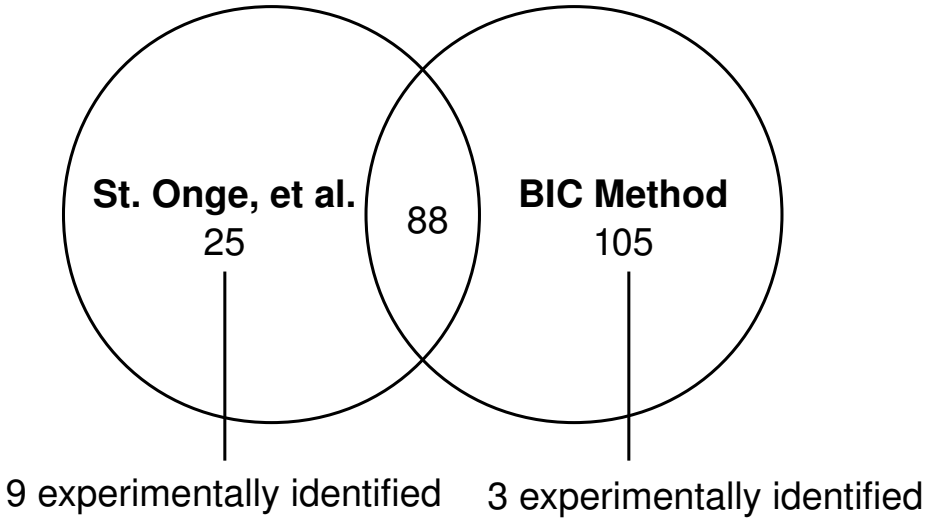

FIGURE S4.—Venn diagram comparing the overlap of epistatic pairs found examining the ST ONGE *et al.* (2007) dataset using the authors' original method and our BIC method (including all three epistatic subtypes and their corresponding null models with the FDR procedure). For the pairs that only one method identified, we indicate the number for which an interaction has already been experimentally identified (obtained using the BIOGRID database).

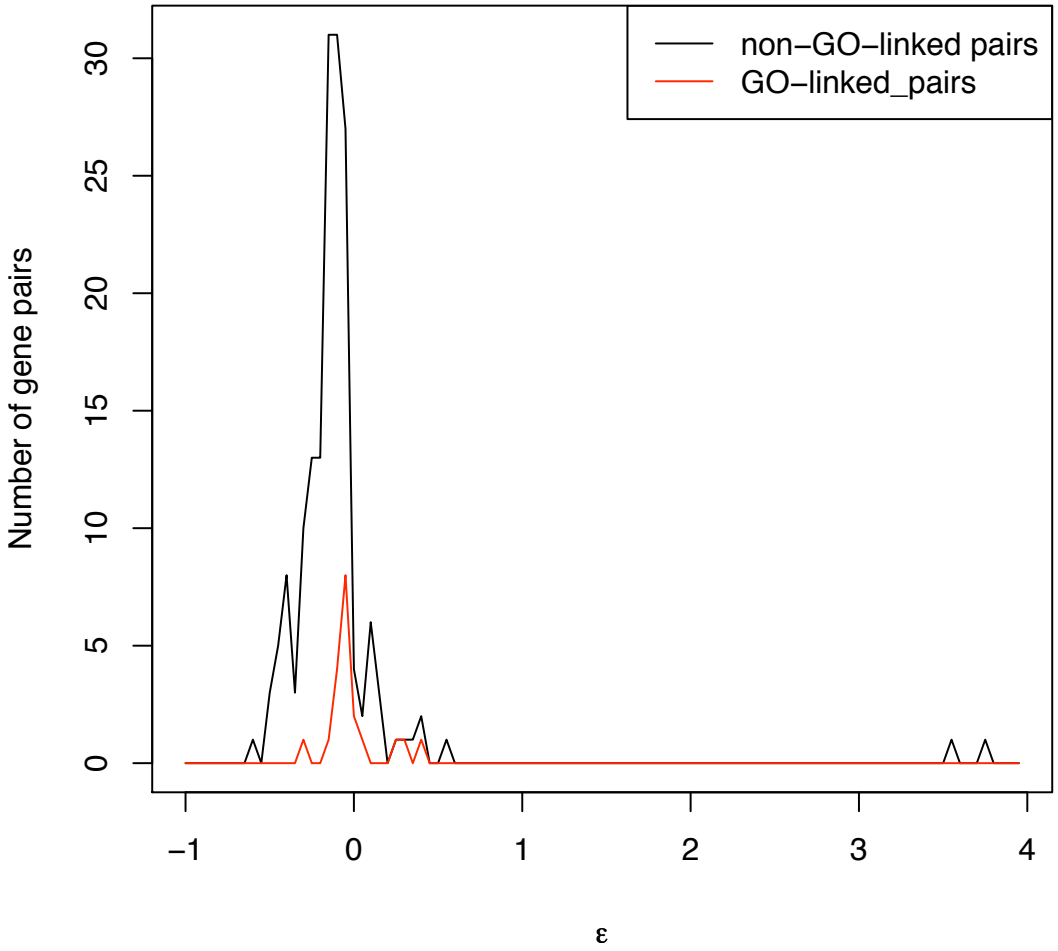

FIGURE S5.—Distribution of the epistasis values (ε) for significant epistatic pairs, determined using the FDR procedure with the BIC method including all three epistatic subtypes and their corresponding null models, for all gene pairs of the ST. ONGE *et al.* (2007) dataset in the presence of MMS (bin size = 0.05). ε values for gene pairs with specific functional links as determined by GO terms (explained in text) are shown in red (mean of -0.060, not significantly different than 0), and for those without specific functional links are shown in black. A similar figure is shown in ST. ONGE *et al.* (2007).

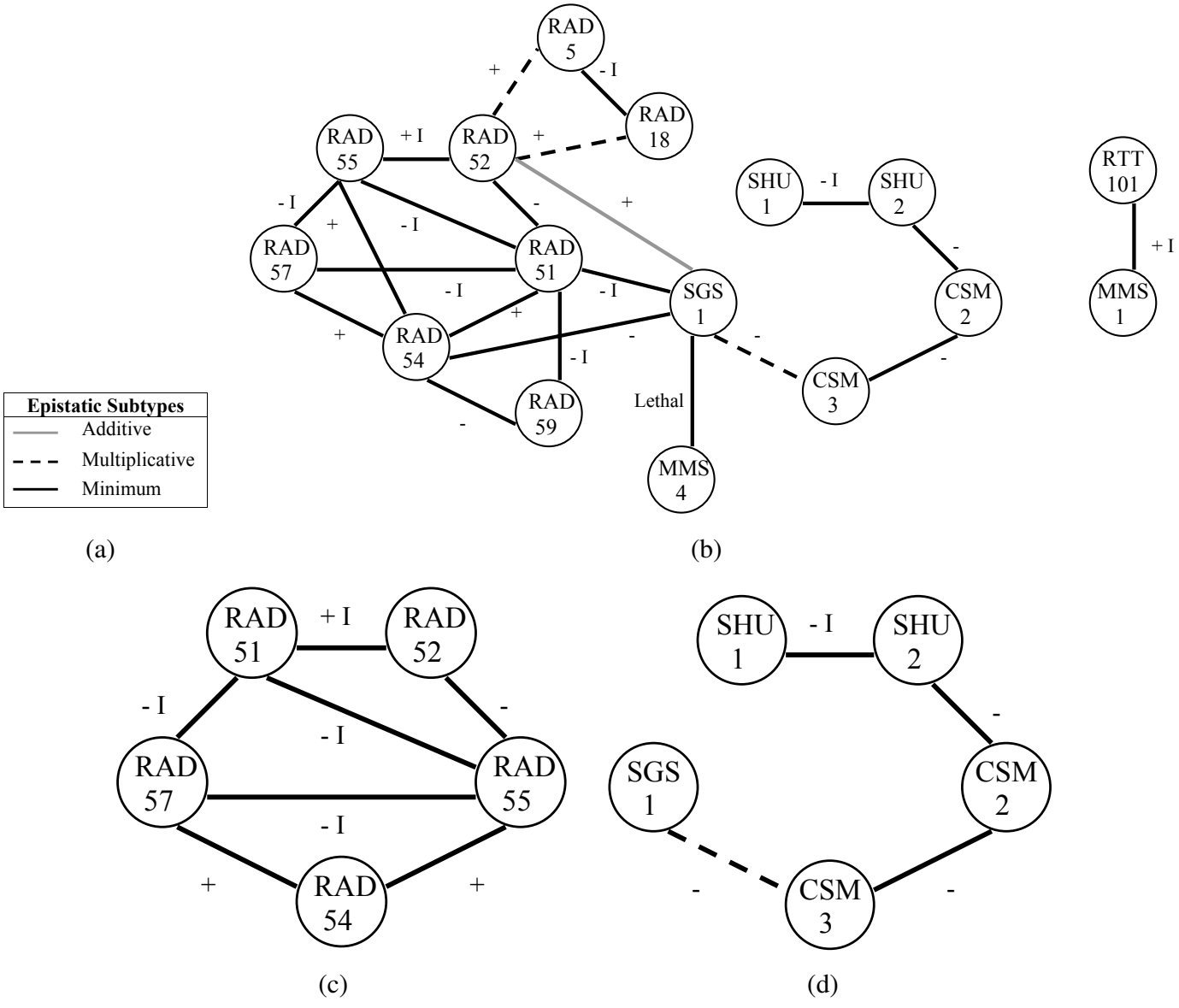

FIGURE S6.—Several epistatic networks constructed based on the ST. ONGE *et al.* (2007) dataset (using the FDR procedure with the BIC method including all three epistatic subtypes and their corresponding null models). Legend in (a) denotes the possible subtypes for each epistatic interaction. (b) depicts all epistatic interactions that were found between genes that also share "specific" functional links (defined in text); note that epistatic interactions were inferred between other genes in this network that do not have a specific functional link (not shown). (c) represents all inferred interactions in the RAD homologous recombination pathway. (d) represents all inferred interactions in the SHU pathway. "+" sign indicates positive epistasis, "-" sign indicates negative epistasis, and "I" indicates that the interaction has been experimentally identified.

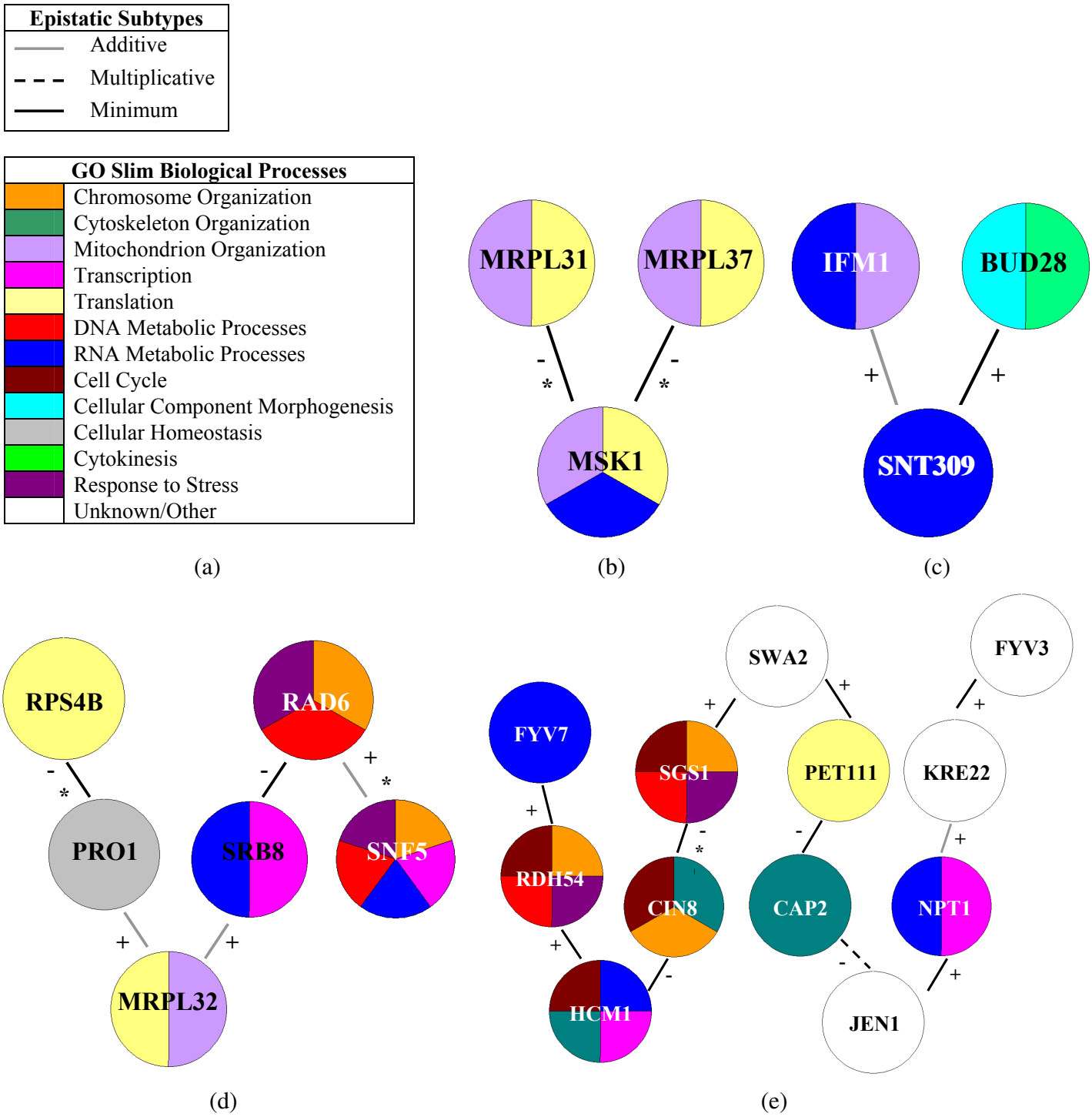

FIGURE S7.—Several epistatic networks constructed based on the JASNOS and KORONA (2007) dataset (using the FDR procedure with the BIC method including all three epistatic subtypes and their corresponding null models). Legend in (a) shows each type of epistatic interaction and each color which corresponds to one biological process as listed by the Gene Ontology (GO Slim terms). (b), (c), (d), and (e) are examples of inferred three-gene, three-gene, six-gene, and twelve-gene networks, respectively. "+" sign indicates positive epistasis, "-" sign indicates negative epistasis, "*" indicates that the gene pair has a specific functional link (described in text).

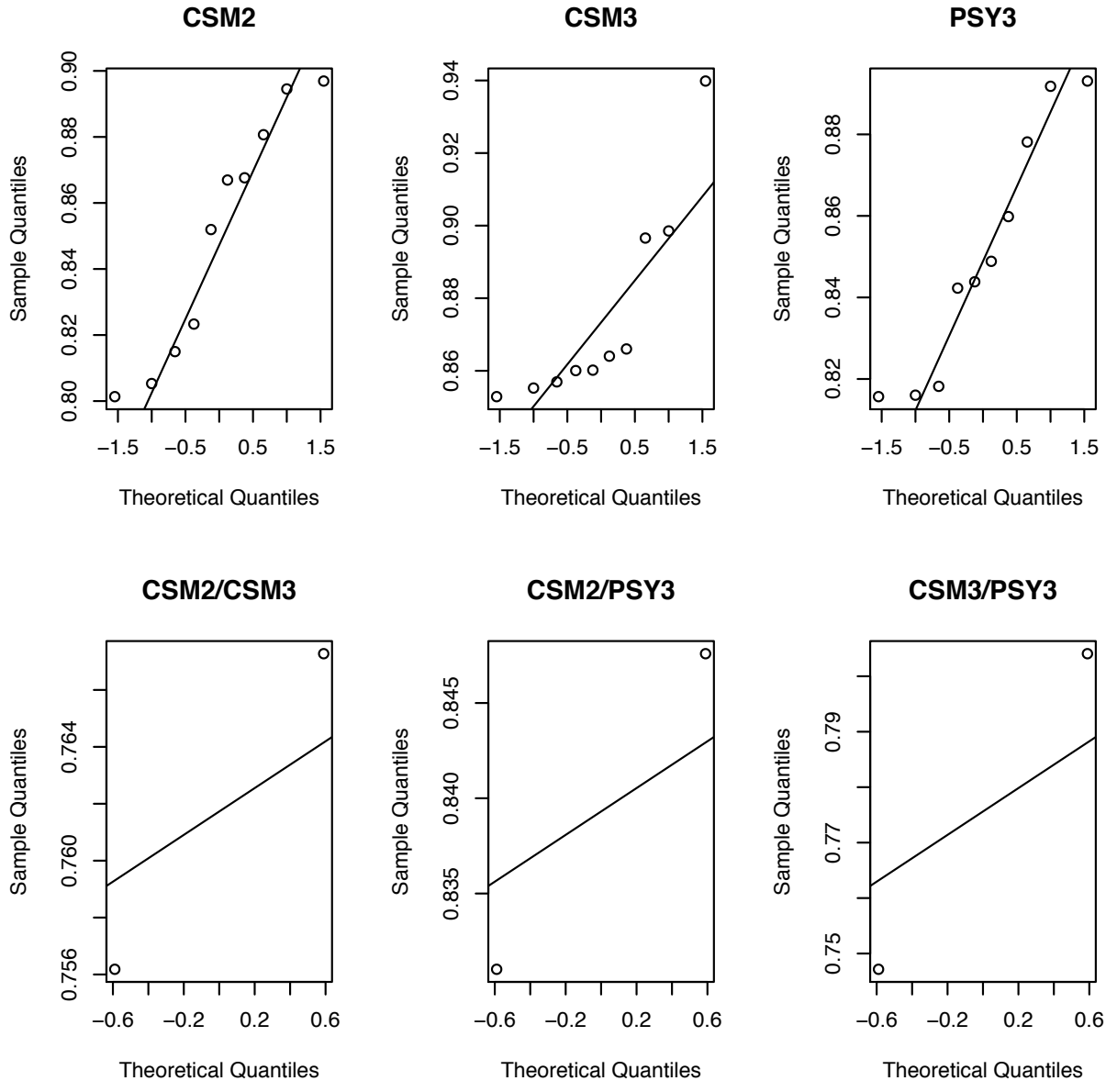

FIGURE S8.—Quantile-Quantile plot of the single-mutant fitness values for the genes CSM2, CSM3, and PSY3 (top row), and the double-mutant fitness values for CSM2/CSM3, CSM2/PSY3, and CSM3/PSY3 (bottom row) (ST. ONGE *et al.* 2007) dataset. We identify an interaction between CSM2 and CSM3, but not for the other two pairs, when analyzing this dataset.

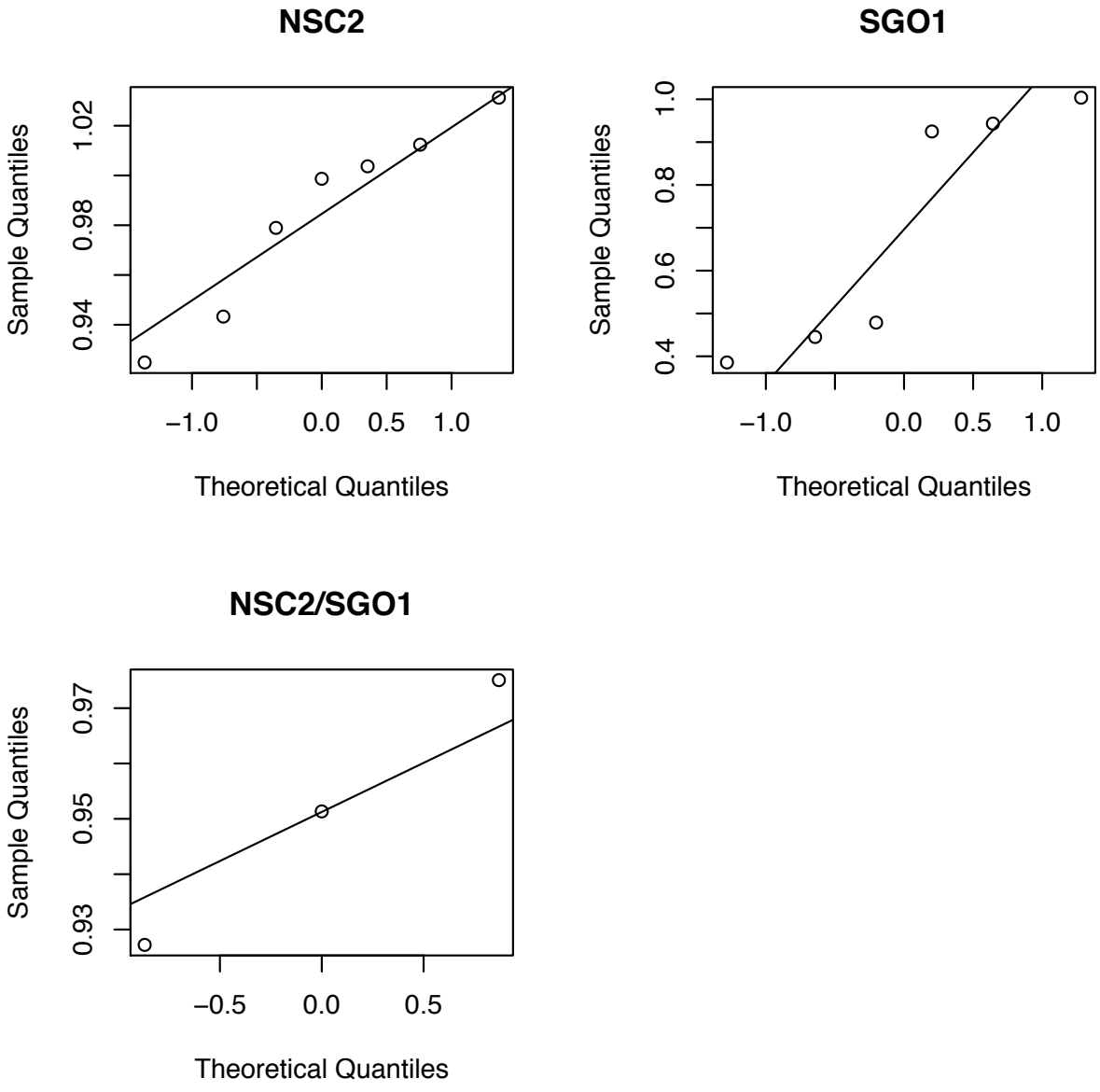

FIGURE S9.—Quantile-Quantile plot of the single-mutant fitness values for the genes NCS2 and SGO1 (top row), and the double-mutant fitness values of NCS2/SGO1 (bottom row) (JASNOS and KORONA (2007) dataset). We do not identify an interaction between these two genes.

TABLE S1

Fractions of simulations that recover the true model among the six models when allowing both beneficial and deleterious mutations, i.e., the range for fitness of single mutants is [0.0, 1.2] and the range for fitness of double mutants is [0.0, 1.5]

μ_i	μ_j	$P_N^{(a)}$	$P_N^{(p)}$	$P_N^{(m)}$	ϵ_{ij}	$P_E^{(a)}$	$P_E^{(p)}$	$P_E^{(m)}$
0.8	0.8	0.987	0.986	0.993	-0.3	0.996	0.903	0.929
0.8	0.8				-0.2	0.991	0.899	0.956
0.8	0.8				-0.1	0.964	0.869	0.982
0.8	0.8				0.1	0.742	0.719	0.997
0.8	0.8				0.2	0.581	0.602	0.998
0.8	0.8				0.3	0.740	0.613	0.998
0.9	0.6	0.989	0.987	0.994	-0.3	0.999	0.964	0.982
0.9	0.6				-0.2	0.997	0.948	0.971
0.9	0.6				-0.1	0.986	0.901	0.979
0.9	0.6				0.1	0.878	0.855	0.997
0.9	0.6				0.2	0.836	0.843	0.999
0.9	0.6				0.3	0.891	0.837	0.999

For all the simulations, the fitness standard deviation $\sigma_i = \sigma_j = \sigma_{ij} = 0.05$. $P_N^{(p)}$ and $P_N^{(m)}$ are defined in a manner similar to $P_N^{(a)}$ for the multiplicative and minimum models, respectively. $P_E^{(a)}$ represents the power of our method, i.e. the percentage of simulations (simulated under the additive epistatic model with epistasis coefficient (ε_{ij}) with the additive epistatic model attaining the minimum value of the BIC among the six models. $P_E^{(p)}$ and $P_E^{(m)}$ are defined in a manner similar to $P_E^{(a)}$ for the multiplicative and minimum models, respectively. 50 replicate fitness values are simulated.

TABLE S2

Fractions of simulations that recover the true model among the six models with different values of the epistasis coefficient (ε_{ij}) and mean fitnesses of single mutants (μ_i) and (μ_j)

μ_i	μ_j	$P_N^{(a)}$	$P_N^{(p)}$	$P_N^{(m)}$	ϵ_{ij}	$P_E^{(a)}$	$P_E^{(p)}$	$P_E^{(m)}$
0.8	0.8	0.261	0.195	0.338	-0.3	0.527	0.315	0.469
0.8	0.8				-0.2	0.519	0.285	0.511
0.8	0.8				-0.1	0.498	0.232	0.533
0.8	0.8				0.1	0.444	0.175	0.591
0.8	0.8				0.2	0.36	0.199	0.599
0.8	0.8				0.3	0.255	0.344	0.603
0.9	0.6	0.273	0.161	0.301	-0.3	0.514	0.341	0.455
0.9	0.6				-0.2	0.509	0.302	0.494
0.9	0.6				-0.1	0.512	0.231	0.505
0.9	0.6				0.1	0.492	0.173	0.567
0.9	0.6				0.2	0.446	0.202	0.557
0.9	0.6				0.3	0.394	0.274	0.555

For all the simulations, the fitness standard deviation $\sigma_i = \sigma_j = \sigma_{ij} = 0.05$ and the number of replicates is equal to 2. The rest of the parameters are defined in the same manner as in Table S1.

TABLE S3

Summary of gene pairs with the indicated epistatic subtypes, inferred using the FDR procedure with the BIC method which considers all three epistatic subtypes and their corresponding null models (ST. ONGE *et al.* (2007)) dataset, with fitnesses measured in the absence of MMS)

Epistatic Subtype	Study S (No MMS)
All	171 (100%) $\bar{\epsilon}$ = -0.063*** $\epsilon_{q_{0.5}}$ = -0.057
Additive	4 (2.3%) $\bar{\epsilon}$ = 0.049* $\epsilon_{q_{0.5}}$ = 0.074
Multiplicative	7 (4.1%) $\bar{\epsilon}$ = -0.109** $\epsilon_{q_{0.5}}$ = -0.084
Minimum	155 (90.6%) $\bar{\epsilon}$ = -0.063*** $\epsilon_{q_{0.5}}$ = -0.057

Numbers are the counts of each type, and percentages are given of the total number of epistatic pairs. The mean ($\bar{\epsilon}$) and median ($\epsilon_{q0.5}$) of the epistatic parameter (ϵ) are given for each subtype, with "*" indicating that the mean of ϵ is significantly different from 0 (p-value ≤ 0.05; "**" implies p-value ≤ 0.01, "***"implies p-value ≤ 0.001). Note that 5 of the epistatic pairs are synthetic lethal (not shown in table).

TABLE S4

Comparison of validation measures for each dataset for different variations of the FDR and BIC procedure, considering only a subset of epistatic subtypes with their corresponding null models: all epistatic subtypes (A, P, M), only the additive and multiplicative subtypes (A, P), only the additive (A), only the multiplicative (P), or only the minimum (M) subtype (see text for details).

	Subtypes Considered in BIC Procedure				
	A, P, M	A, P	A	P	M
Study S (No MMS)					
Number Found (323)	171	120	156	99	203
Functional Links (36)	20 (0.4391)	20 (0.0135)*	25 (0.0056)*	16 (0.0460)*	18 (0.9682)
GO Terms (Biological Process) (283)	158 (0.0058)*	114 (0.0010)*	145 (0.0037)*	95 (0.0011)*	186 (0.0043)*
GO Terms (All) (307)	165 (0.1558)	115 (0.4151)	149 (0.4548)	95 (0.4241)	196 (0.08893)
Experimentally Identified (29)	14	14	18	12	14

Numbers in parentheses indicate p-values by Fisher's Exact Test, ``*" indicates significance. Study S refers to the ST. ONGE *et al.* (2007) dataset, measured in the absence of MMS. Numbers in parentheses indicate the total number of tested pairs and the total number of each type of link found in each complete dataset.

TABLE S5

Comparison of validation measures for each dataset for different variations of the FDR and BIC procedure, considering only a subset of epistatic subtypes with their corresponding null models: all epistatic subtypes (A, P, M), only the additive and multiplicative subtypes (A, P), only the additive (A), only the multiplicative (P), or only the minimum (M) subtype (see text for details).

	Subtypes Considered in BIC Procedure				
	A, P, M	A, P	A	P	M
Study J					
Number Found (636)	352	273	263	231	329
Functional Links (25)	19 (0.0240)*	13 (0.2274)	11 (0.4700)	10 (0.4216)	15 (0.2466)
GO Terms (Biological Process) (115)	69 (0.2720)	50 (0.6226)	55 (0.1328)	44 (0.4726)	68 (0.0)*
GO Terms (All) (369)	224 (0.0220)*	172 (0.1180)	160 (0.4112)	146 (0.1172)	213 (0.007)*
Experimentally Identified (3)	3	2	1	2	3
Study S (MMS)					
Number Found (323)	193	192	247	171	243
Functional Links (36)	21 (0.473)	29 (0.0004)*	34 (0.0)*	29 (0.0)*	24 (0.8102)
GO Terms (Biological Process) (283)	174 (0.0652)	174 (0.0368)*	223 (0.0108)*	153 (0.1822)	213 (0.5560)
GO Terms (All) (307)	185 (0.2806)	182 (0.6940)	237 (0.1458)	162 (0.7006)	231 (0.5794)
Experimentally Identified (29)	14	22	24	23	21
Study S (No MMS)					
Number Found (323)	171	120	156	99	203
Functional Links (36)	20 (0.2986)	20 (0.0052)*	25 (0.0018)*	16 (0.0258)*	18 (0.9198)
GO Terms (Biological Process) (283)	158 (0.0052)*	114 (0.0008)*	145 (0.0036)*	95 (0.001)*	186 (0.004)*
GO Terms (All) (307)	165 (0.1520)	115 (0.4182)	149 (0.4510)	95 (0.4264)	196 (0.088)
Experimentally Identified (29)	14	14	18	12	14

Numbers in parentheses indicate p-values by permutation, "*" indicates significance. Study J refers to the JASNOS and KORONA (2007) dataset, and Study S refers to the ST. ONGE *et al.* (2007) dataset. Numbers in parentheses indicate the total number of tested pairs and the total number of each type of link found in each complete dataset.

TABLE S6

Fractions of simulations that recover the true model among the six models, with different values of the epistasis coefficient (ϵ_{ij}) and mean fitnesses of single mutants (μ_i) and (μ_j)

μ_i	μ_j	$P_N^{(a)}$	$P_N^{(p)}$	$P_N^{(m)}$	ϵ_{ij}	$P_E^{(a)}$	$P_E^{(p)}$	$P_E^{(m)}$
0.8	0.8	0.842	0.824	0.926	-0.3	0.836	0.574	0.705
0.8	0.8				-0.2	0.807	0.533	0.732
0.8	0.8				-0.1	0.758	0.396	0.762
0.8	0.8				0.1	0.59	0.423	0.871
0.8	0.8				0.2	0.472	0.335	0.875
0.8	0.8				0.3	0.491	0.506	0.898
0.9	0.6	0.856	0.821	0.903	-0.3	0.85	0.629	0.742
0.9	0.6				-0.2	0.833	0.589	0.753
0.9	0.6				-0.1	0.796	0.405	0.732
0.9	0.6				0.1	0.677	0.374	0.848
0.9	0.6				0.2	0.64	0.487	0.873
0.9	0.6				0.3	0.645	0.527	0.884

For all simulations, the fitness standard deviation $\sigma_i = \sigma_j = \sigma_{ij} = 0.05$ and the number of replicates is equal to 10. The rest of the parameters are defined in the same manner as in Table 1, main text.