Genetic Interactions Between Transcription Factors Cause Natural Variation in Yeast
Justin Gerke, et al.
*Science* **323**, 498 (2009);
DOI: 10.1126/science.1166426

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of February 24, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/cgi/content/full/323/5913/498

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/323/5913/498/DC1

This article cites 43 articles, 11 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/323/5913/498#otherarticles

This article has been cited by 8 article(s) on the ISI Web of Science.

This article has been cited by 3 articles hosted by HighWire Press; see:
http://www.sciencemag.org/cgi/content/full/323/5913/498#otherarticles

This article appears in the following subject collections:
Evolution
http://www.sciencemag.org/cgi/collection/evolution

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2009 by the American Association for the Advancement of Science; all rights reserved. The title Science is a registered trademark of AAAS.
Genetic Interactions Between Transcription Factors Cause Natural Variation in Yeast

Justin Gerke, Kim Lorenz, Barak Cohen*

Our understanding of the genetic basis of phenotypic diversity is limited by the paucity of examples in which multiple, interacting loci have been identified. We show that natural variation in the efficiency of sporulation, the program in yeast that initiates the sexual phase of the life cycle, between oak tree and vineyard strains is due to allelic variation between four nucleotide changes in three transcription factors: **IME1**, **RME1**, and **RSF1**. Furthermore, we identified that selection has shaped quantitative variation in yeast sporulation between strains. These results illustrate how genetic interactions between transcription factors are a major source of phenotypic diversity within species.

Understanding the molecular basis of natural phenotypic diversity is a major challenge in modern genetics (1–6). Knowing how individual genetic polymorphisms combine to produce phenotypic change could strengthen evolutionary theory and advance applications such as personalized medicine (7, 8). Many loci that contribute to variation have been identified across taxa, but only a small fraction has been resolved to the nucleotide level (9, 10). Examples of complex traits in which causative polymorphisms have been identified at multiple contributing loci are even rarer (11). As a result, the interactions between nucleotide changes in nature, and thus the genetic mechanisms of phenotypic change, are largely unknown.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Nearest marker</th>
<th>lod score</th>
<th>Variance explained (%)</th>
<th>Additive effect (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>L7.9</td>
<td>86.42</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>L7.17</td>
<td>4.7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>L10.14</td>
<td>68.2</td>
<td>29</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>L11.2</td>
<td>3.9</td>
<td>1</td>
<td>–3</td>
</tr>
<tr>
<td>13</td>
<td>L13.6</td>
<td>28.7</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

*To whom correspondence should be addressed. E-mail: cohen@genetics.wustl.edu

Crosses of laboratory strains of the yeast **Saccharomyces cerevisiae** have identified genes and polymorphisms governing complex traits (12–21). However, these lines harbor laboratory-engineered gene deletions and deleterious mutations that are pleiotropic for multiple traits, which may obscure the natural genetic architecture. The natural diversity of this species, which includes isolates from clinical, vineyard, and oak tree environments, remains largely untapped (22–24).

Sporulation efficiency is a highly heritable complex trait that varies among natural populations of **S. cerevisiae** (25). Sporulation is a core developmental program that initiates the sexual phase of the yeast life cycle and promotes long-term survival during dessication or starvation (26). Sporulation is triggered as a response to environmental change (27) and is hypothesized to be under different selective pressures in different habitats (28). Accordingly, wild isolates from North American oak trees and associated soil samples sporulate with efficiencies approaching 100%, but strains isolated from naturally occurring vineyard fermentations sporulate at lower efficiencies.
rates (25). Genetic dissection of the underlying natural variation in sporulation efficiency provides an opportunity to uncover the nucleotide changes that govern ecologically driven variation.

To identify quantitative trait loci (QTL) governing natural variation in sporulation, we crossed YPS606, a strain isolated from the bark of an oak tree in Pennsylvania that sporulates at 99% efficiency, and BC187, a strain originating from a California wine barrel that sporulates at only 3.5% (25). We developed genetic markers by shotgun sequencing these parent strains and produced a genetic linkage map of 225 loci typed in 374 recombinant segregants. This map covers the yeast genome at an average of 11 centimorgan intervals (table S1). Composite interval mapping (29) was used to identify five QTL on four chromosomes (fig. S1 and Table 1) that significantly cosegregated with variation in sporulation efficiency with an experiment-wide error rate of P = 0.05 [logarithm of the odds ratio for linkage (lod) score > 3.1, which is significant by permutation analysis]. Alleles from the oak parent at four of the QTL were linked to an increase in sporulation efficiency (markers L7.9, L7.17, L10.14, and L13.6). At the locus on chromosome 11 (marker L11.2), an allele from the poorly sporulating vineyard parent promotes more efficient sporulation in segregating progeny. The presence of a vineyard allele promoting higher sporulation was on the basis of previously observed transgressive segregants that sporulate more efficiently than the oak parent (25).

We quantified the total amount of variation explained by these five QTL using a linear model trained on 200 segregants (table S2) (29). In an independent test set of 155 segregants, the model explains 88% of the phenotypic variation [squared correlation coefficient (R²) = 0.88], with an average prediction error of ±8% (Fig. 1). This indicates that the five QTL explain most of the variation in sporulation efficiency in this cross. To obtain the best fit, the model must incorporate two- and three-way interactions between loci (F-test, P < 0.001) (29), indicating that the genetic architecture of sporulation efficiency is nonadditive and complex. Although incorporating the effects of all five loci produces the best-fit model (F-test, P < 0.001), only three of the loci have large effects. When the minor QTL (markers L7.17 and L11.2) are ignored, the results are virtually identical (R² = 0.87, prediction error ±8%).

One major QTL (marker L7.9) covers a 100-kb confidence interval on chromosome 7. RME1, a transcription factor that suppresses sporulation in specific cell types (30), resides in this peak (fig. S1A). To test whether allelic variation in RME1 produces variation in sporulation efficiency, we deleted each parental allele of RME1 in a hybrid background [through reciprocal hemizygosity analysis (18)]. This showed that the allelic contributions of RME1 from each parent were different, confirming that variation in RME1 affects sporulation efficiency (Fig. 2A). The coding region of RME1 contains no amino acid substitutions between the oak and vineyard parents, which suggests that the allelic difference is regulatory. By replacement (29), we confirmed that a single nucleotide insertion/deletion 308 base pairs (bp) upstream of the initiation codon (fig. S2A) accounts for the effect of the RME1 locus on sporulation.

Fig. 1. A linear model of the effects of five QTL on sporulation efficiency. Expected values of sporulation efficiency plotted as a function of observed values for 155 segregants. Expected values are derived from a linear model based on 200 independent segregants (table S2).

Fig. 2. (A) Reciprocal hemizygosity analysis of RME1, IME1, and RSF1 in an oak/vineyard hybrid background. O, oak parent allele; V, vineyard parent allele. (B) Sporulation efficiency in oak, in oak with the vineyard parent allele [RME1(del-308A)], and in oak with the entire vineyard parent locus, including all the coding and noncoding polymorphisms. Variation among multiple replicate clones of the experiment (F test, P < 0.001) explains the discrepancy between the RME1(del-308A) and full allele in the oak background, and this is not an effect of the two replacement types (F test, P = 0.13). (C) Replacement of both the oak IME1 causative nucleotides with the vineyard parent alleles (L325M and A-548G) in the oak parent is equivalent to replacing the entire IME1 vineyard parent locus, coding and noncoding, in the oak parent (full locus). (D) Replacements of all four alleles in the vineyard parent strain (oak alleles) compared with placing the vineyard alleles in the oak parent. Error bars, ±1 SD.
ulation efficiency (Fig. 2B). The vineyard strain allele [RME1(del-308A)], which has a deletion of a single adenine relative to the oak strain, also reduces sporulation efficiency in laboratory strains (21). This nucleotide change presumably increases the expression of RME1, which represses sporulation, as the RME1(del-308A) allele is expressed at higher levels than the oak allele (25, 31).

A second major QTL (L10.14) located in a 50-kb confidence interval on chromosome 10 also contained a strong candidate gene, IME1 (fig. S1B). IME1 is a transcriptional activator and master regulator that initiates yeast sporulation (32). Reciprocal hemizygosity analysis confirmed that IME1 quantitatively controls sporulation efficiency (Fig. 2A). We identified 8 polymorphisms in the coding region, both synonymous and nonsynonymous, and 39 polymorphisms in the noncoding regions of IME1 between the oak and vineyard strain. Allele replacements demonstrated that two of these polymorphisms account for the full effect of the IME1 locus on sporulation efficiency (Fig. 2C). Although other polymorphisms in this region may affect sporulation efficiency, they must be redundant with the two causative alleles that we identified. We identified a causative nonsynonymous substitution in the vineyard strain, IME1 (L325M) (fig. S2B), which resides in an IME1 domain essential for protein-protein interactions with Rim11 and Ume6, which also regulate the initiation of sporulation (33). Mutation of this leucine reduces the ability of IMe1 to activate transcription (34). We also identified a noncoding IME1(A-548G) polymorphism in an 11-bp sequence that is conserved among three yeast species closely related to S. cerevisiae (fig. S2C).

The third major QTL affecting sporulation occurs in a 100-kb region of chromosome 13 (L13.6, fig. S1D). Reciprocal hemizygosity analysis of genes in this region identified RSF1 as a candidate gene affecting sporulation efficiency (Fig. 2A). Allele replacements (fig. S3) confirmed that a single derived polymorphism in the vineyard strain—coding for a substitution of a conserved glutamic acid with a glycine RSF1(D181G)—is responsible for the allelic effect of RSF1 (fig. S2D). RSF1 encodes a transcriptional activator of mitochondrial genes critical for cellular respiration (35). Because a respiratory signal promotes IME1 expression and sporulation (36), the vineyard allele likely reduces the function of Rsf1.

In total, we identified four nucleotide changes causing variation in sporulation efficiency by directly affecting three transcription factors governing pathways that regulate the initiation of sporulation. Our QTL model (table S2) includes, however, potentially undiscovered linked alleles that could account for some of the variation. Therefore, we engineered yeast strains isogenic to each parent but carrying the causative alleles from the opposite parent (Fig. 2D) (29). The vineyard parental strain, which sporulates at 3.5 ± 0.1%, increases to 78 ± 2% when carrying all four oak alleles. In the oak parent, the replacement of these four nucleotides reduced sporulation efficiency from 99 ± 0.2% to 14.9 ± 1% and places the phenotype of the oak strain background within the range normally seen only among vineyard strains (Fig. 3).

Our QTL model also predicts that the four causative nucleotides will interact. We therefore compared the phenotypes of strains isogenic to the oak background carrying all possible permutations of the four oak and vineyard alleles. We chose the oak strain background for this experiment because isolates of other species of yeast sporulate efficiently, supporting the idea that the oak strain resembles the ancestral state of sporulation efficiency from which we hypothesize the vineyard parent alleles arose. This is further supported by the fact that three of the four vineyard alleles reducing sporulation [IME1(L325M), IME1(A-548G), and RSF1(D181G)] are derived (fig. S2).

Analysis of the allele replacement strains revealed extensive interactions among the four nucleotides as all possible two-, three-, and four-way interactions are statistically significant in an analysis of variance (table S3). The interactions indicate that the vineyard alleles work synergistically to reduce sporulation efficiency. For example, the IME1 coding and noncoding polymorphisms interact and double the sum of their individual effects (Fig. 4A). The strongest interactions observed occurred between RME1(del-308A) and the two polymorphisms at IME1 (Fig. 4B). The
RSF1 vineyard allele, which causes less than a 3% change on its own, was also found to interact synergistically with the other vineyard alleles (Fig. 4C) and was responsible for 11%, 13%, and 16% drops in sporulation efficiency through its two-way interactions (table S3). These results illustrate mechanisms by which combinations of alleles can produce a phenotypic change larger than expected from individual effects.

Genetic interactions (epistasis) are often seen between genomic regions affecting quantitative traits (37), and this study demonstrates how a small number of nucleotides can create complex, quantitative variation in phenotype highlighting the importance of single nucleotides on epistasis. This emphasizes the need to incorporate genetic interactions into models that seek to accurately predict phenotype from genotype. If prevalent, genetic interactions between nucleotides will be a major hurdle in the endeavor to connect genetic and phenotypic variation in humans (38). Our identification of epistasis between RME1 and IME1 identifies the idea that a search for epistasis should incorporate previous knowledge of functional relationships between genes and proteins (39).

We found that alleles reducing sporulation efficiency were at varying frequencies in additional vineyard strains (table S4). The RME1 polymorphism was found to be ubiquitous, whereas the IME1 polymorphisms were the rarest. More than half of the vineyard strains harbor at least two of the alleles reducing sporulation, which suggests that the two-way interactions we identified may be producing phenotypic change in nature. However, the four nucleotides we identified are not sufficient to predict all of the differences in sporulation efficiency among vineyard strains. For example, strain UCDS1 sporulates less efficiently than M34 and M15 (Fig. 3), despite carrying fewer of the alleles reducing sporulation efficiency than M34 and M15 (Fig. 3), de-