Regulatory Rewiring in a Cross Causes Extensive Genetic Heterogeneity

Takeshi Matsui, Robert Linder, Joann Phan, Fabian Seidl, and Ian M. Ehrenreich[1]
Molecular and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-2910

ABSTRACT Genetic heterogeneity occurs when individuals express similar phenotypes as a result of different underlying mechanisms. Although such heterogeneity is known to be a potential source of unexplained heritability in genetic mapping studies, its prevalence and molecular basis are not fully understood. Here we show that substantial genetic heterogeneity underlies a model phenotype—the ability to grow invasively—in a cross of two *Saccharomyces cerevisiae* strains. The heterogeneous basis of this trait across genotypes and environments makes it difficult to detect causal loci with standard genetic mapping techniques. However, using selective genotyping in the original cross, as well as in targeted backcrosses, we detected four loci that contribute to differences in the ability to grow invasively. Identification of causal genes at these loci suggests that they act by changing the underlying regulatory architecture of invasion. We verified this point by deleting many of the known transcriptional activators of invasion, as well as the gene encoding the cell surface protein Flo11 from five relevant segregants and showing that these individuals differ in the genes they require for invasion. Our work illustrates the extensive genetic heterogeneity that can underlie a trait and suggests that regulatory rewiring is a basic mechanism that gives rise to this heterogeneity.

KEYWORDS complex traits; genetic mapping; invasive growth; regulatory networks; yeast

GENETIC studies in humans and model organisms have reported unexplained heritability for many traits (Manolio *et al.* 2009). A possible contributor to this "missing" heritability is genetic heterogeneity—individuals exhibiting similar phenotypes owing to different genetic and molecular mechanisms (Risch 2000; McClellan and King 2010; Wray and Maier 2014). Genetic heterogeneity can reduce the statistical power of mapping studies (Manchia *et al.* 2013; Wray and Maier 2014) and may involve multiple variants segregating in the same gene (*allelic* heterogeneity) or different genes (*nonallelic* heterogeneity) (Risch 2000). Work to date has shown that allelic heterogeneity is widespread (*e.g.*, McClellan and King 2010; Ehrenreich *et al.* 2012; Long *et al.* 2014) and often involves two or more null or partial loss-of-function variants segregating in a single phenotypically important gene (*e.g.*, Nogee *et al.* 2000; Sutcliffe *et al.* 2005; Will *et al.* 2010). However, the prominence and underlying mechanisms of nonallelic heterogeneity are less understood.

In this paper we describe an example of nonallelic heterogeneity using heritable variation in the ability of *Saccharomyces cerevisiae* strains to undergo haploid invasive growth as our model. Invasive growth is a phenotype that is triggered by low carbon or nitrogen availability and is thought to be an adaptive response that allows yeast cells to adhere to and penetrate surfaces (Cullen and Sprague 2000). Invasion typically requires expression of *FLO11*, which encodes a cell surface glycoprotein that facilitates cell-cell and cell-surface adhesion (Lo and Dranginis 1998; Rupp *et al.* 1999). In addition to *FLO11*, *S. cerevisiae* possesses other cell surface proteins that can contribute to adhesion-related traits [as described in Guo *et al.* (2000) and Halme *et al.* (2004) and elsewhere]. In some cases, these cell surface proteins are regulated by multiple signaling cascades (Bruckner and Mosch 2012), potentially providing an opportunity for genetic variants in different pathways to have similar effects on invasion.

Here we examine the genetic basis of variation in the ability to invade on two carbon sources—glucose and ethanol—in

Copyright © 2015 by the Genetics Society of America
doi: 10.1534/genetics.115.180661
Manuscript received March 12, 2015; accepted for publication July 28, 2015; published Early Online July 30, 2015.
Supporting information is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.180661/-/DC1
[1]Corresponding author: Molecular and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA 21546. E-mail: ian.ehrenreich@usc.edu

a cross of the laboratory strain BY4716 (BY) and the clinical isolate YJM789 (YJM) (Liti *et al.* 2009). YJM is highly invasive on both carbon sources (Figure 1A). In contrast, BY cannot grow invasively on either carbon source (Figure 1A). This is because BY carries a nonsense allele of *FLO8* (Figure 1B; see also *Materials and Methods*), which encodes a transcriptional activator that is regulated by the Ras-cAMP-PKA pathway. Flo8 is typically required for invasive growth in both *S. cerevisiae* (Liu *et al.* 1996) and *Candida albicans* (Cao *et al.* 2006). Consistent with the importance of *FLO8* for invasion, deletion of this gene from YJM significantly reduces its invasive growth on both carbon sources (Figure 1B; see also *Materials and Methods*).

While screening BYxYJM segregants for invasion on the two carbon sources, we found that many individuals exhibit invasion even though they possess the *FLO8*^{BY} nonsense allele, a result that also was recently reported by Song *et al.* (2014). We show that this *FLO8*-independent growth has a heterogeneous genetic basis that reflects the presence of multiple distinct regulatory architectures that enable *FLO8*-independent invasion. Most of these regulatory architectures are *FLO11* dependent but require different transcriptional activators; however, we also provide evidence for an architecture that is *FLO11* independent. Our results suggest that regulatory rewiring is an important source of nonallelic genetic heterogeneity and illustrate how studying the causes of phenotypic similarities among genetically distinct individuals can advance our understanding of complex traits.

Materials and Methods

Generation of initial mapping population

We used the synthetic genetic array marker system (Tong *et al.* 2001) to generate recombinant BYxYJM *MAT***a** segregants. The BY parent of our cross was *MAT*α *can1*Δ::*STE2pr-SpHIS5 lyp1*Δ *his3*Δ, while the YJM parent was *MAT***a** *his3*Δ::*natMX ho*::*kanMX*. We mated these BY and YJM haploids to produce the diploid progenitor of our cross, which was sporulated using standard techniques (Sherman 1991). *MAT***a** segregants were obtained using random spore plating on minimal medium containing canavanine, as described previously (Ehrenreich *et al.* 2010; Taylor and Ehrenreich 2014).

Phenotyping for invasive growth

Strains were phenotyped for invasive growth on 2% agar plates containing yeast extract and peptone (YP) with either 2% glucose (dextrose) or 2% ethanol as the carbon source (YPD and YPE, respectively). Prior to pinning onto the agar plates, strains were grown overnight to stationary phase in liquid YPD. After this culturing step, strains were then pinned onto agar plates and allowed to grow for 5 days. Following this incubation period, we screened for invasive growth by applying water to the agar plates, manually scrubbing colonies, and decanting the mixture of water and cells. Presence or absence of invasion was scored by eye under a light microscope. Each segregant was phenotyped three independent times, and the median phenotype was used in analyses (Supporting Information, Table S1).

Genotyping by sequencing

Segregants were genotyped by Illumina sequencing. Whole genome libraries were constructed using the Illumina Nextera XT DNA Library Preparation Kit. These libraries then were sequenced in multiplex to at least five times genomic coverage on either an Illumina HiSeq 2000 or an Illumina NextSeq 500 with 100 basepair (bp) $\times$ 100 bp reads. We also sequenced BY and YJM to $\sim$100 times genomic coverage and used the data to identify 57,402 high-confidence SNPs. Reads for segregants were mapped to the BY genome using a Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009) and SAMtools (Li *et al.* 2009). We called genotypes for each individual by taking the base calls at the SNPs and employing a hidden Markov model by chromosome using the HMM package in R, as described by Taylor and Ehrenreich (2014).

Data availability

The sequence data from our experiments is available from the NCBI Sequence Read Archive under accession numbers SRR2039809–SRR2039935, SRR2039936–SRR2039992, SRR2040045–SRR2040076, SRR2040023–SRR2040044, and SRR2039993–SRR2040022 (Table S1, Table S2, Table S3, Table S4, Table S5, and Table S6). All other data from the paper are provided in the Supplement or are available by request from the authors.

Detection of loci influencing ability to invade

Allele frequency analyses were computed using the genotype data of all individuals from a particular mapping population that exhibited the same phenotype. To determine the intervals of the identified causal loci, we identified regions where the alleles were either fixed or at a frequency of 95% or higher.

Genetic engineering

Knockouts were generated by PCR amplifying the CORE cassette with homology-tailed primers and then selecting for transformants on G418 (Storici *et al.* 2001). NEB Phusion high-fidelity DNA polymerase was used for PCR under the recommended reaction conditions with 35 cycles and an extension time of 30 s per kilobase. The entire coding region of target genes was deleted in these strains. Correct integration of the CORE cassette was checked for each deletion strain using PCR. Allele replacement strains were constructed using the cotransformation of two partially overlapping PCR products (Figure S1), similar to the work of Erdeniz *et al.* (1997). One product contained the promoter and coding region of the gene to be replaced, while the other included (in order) 60 bp of overlap with the 3′ end of the gene PCR product, *kanMX* or *natMX*, and 30–50 bp of the genomic region immediately downstream of the transcribed portion of the gene. Replacement of a gene was verified using Sanger sequencing.

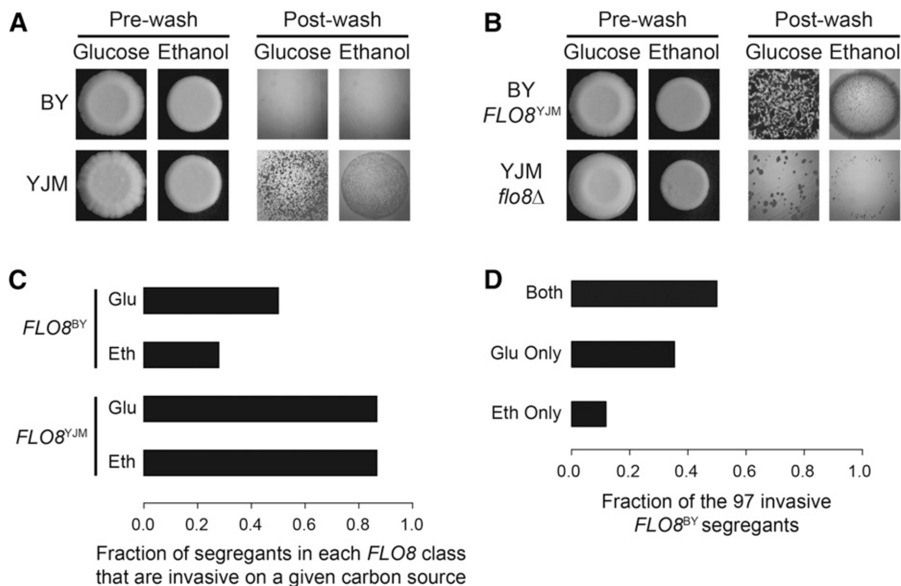

Figure 1 Effects of *FLO8* on ability to invade. (A) BY and YJM were grown for 5 days on YPD or YPE plates at 30°. Colonies were then washed off the plates using water and examined for invasion. (B) Comparison of BY with a functional allele of *FLO8* and YJM *flo8Δ*. (**C**) Fraction of the initial mapping population of 127 F_2 BYxYJM segregants that shows invasion on glucose ("glu") or ethanol ("eth") in each *FLO8* genotype class. (D) Fraction of the 97 invasive *FLO8*^{BY} segregants that shows invasion on glucose, ethanol, or both carbon sources ("both").

Generation of backcross segregants

Backcrosses were conducted by mating a BYxYJM segregant to a *MAT*α *his3Δ* version of BY or YJM. Sporulation and selection for *MAT***a** backcross segregants were performed as described for the initial mapping population.

Screening for mating type and nongenetic effects

To induce mating-type switching in our *MAT***a** segregants, we first deleted *URA3* from these individuals using the *hphMX* cassette with homology-tailed primers, as described earlier. Correct integration of the cassette was verified using PCR and further checked by plating the *ura3Δ* strains onto 5-FOA plates. Next, mating-type switching was performed using the pGAL-*HO* plasmid, as described previously (Herskowitz and Jensen 1991). Otherwise, isogenic *MAT***a** and *MAT*α individuals were mated to produce homozygous diploids. These individuals were sporulated as described earlier, and standard microdissection techniques were used to obtain spores from the homozygous diploids. Tetrads from which all four spores were recovered were then grown on glucose and ethanol and checked for the ability to invade (Table S7).

Amplification of the FLO11 coding region

The entire *FLO11* coding region was PCR amplified using 5′-GGAAGAGCGAGTAGCAACCA as the forward primer and 5′-TTGTAGGCCTCAAAAATCCA as the reverse primer. The sizes of the BY and YJM alleles were compared on a 2% agarose gel.

Results

Many BYxYJM segregants show invasion that is independent of FLO8

We examined a population of 127 genotyped BYxYJM *MAT***a** segregants for the ability to invade on two carbon sources—glucose and ethanol (see *Materials and Methods*). Despite the major role of *FLO8* in the invasion phenotypes of BY and YJM (Figure 1, A and B), we unexpectedly found that a large fraction (52%) of segregants with the *FLO8*^{BY} nonsense allele were capable of invading in at least one condition (Figure 1C). A possible explanation for these individuals' phenotypes is that *FLO8*^{BY} is partially functional in some genetic backgrounds. Flo8 is comprised of a LisH domain (amino acids 72–105) that is involved in physical interactions with the transcription factor Mss11 and a transcriptional activation domain (amino acids 701–799) that is necessary for DNA binding (Kim *et al.* 2014). The nonsense polymorphism in *FLO8*^{BY} occurs after the LisH domain at amino acid 142, suggesting that the truncated Flo8 may retain some functionality. We tested for partial functionality of *FLO8*^{BY} by deleting the entire coding portion of *FLO8* from multiple invasive *FLO8*^{BY} segregants and phenotyping them for invasive growth on glucose and ethanol (see *Materials and Methods*). Complete deletion of *FLO8* had no effect on invasion, suggesting that other mechanisms enable these individuals to grow invasively.

Initial effort to identify loci underlying FLO8-independent invasion

As a first step in identifying the genetic basis of *FLO8*-independent invasion, we screened 384 additional F_2 segregants for invasion on glucose and ethanol. We obtained 55 invasive *FLO8*^{BY} individuals from this experiment, bringing the total number of invasive *FLO8*^{BY} individuals to 97. Among these 97 individuals, 50% were invasive on both glucose and ethanol, 37% were invasive only on glucose, and 12% were invasive only on ethanol (Figure 1D). We genotyped the 55 new individuals using low-coverage genome sequencing and attempted to detect enriched alleles among the larger set of 97 genotyped *FLO8*^{BY} strains that were capable of invasion (see *Materials and Methods*). Although our past work

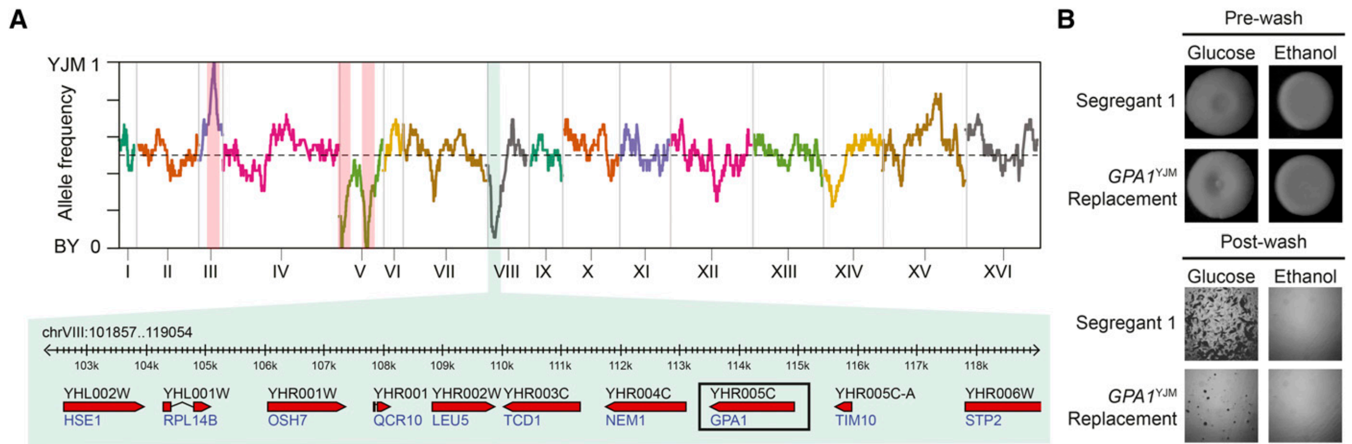

Figure 2 Genetic dissection of *FLO8*-independent glucose-only invasion. (A) Genome-wide relative allele frequency plot of glucose-only *FLO8*^{BY} BYxYJM segregants. *FLO8* and the markers used to generate haploid progeny are highlighted by red vertical bars, while the strongly enriched locus on chromosome VIII, which was nearly fixed for the BY allele, is highlighted by a green vertical bar. The genomic interval underlying the chromosome VIII peak is also provided. (B) Comparison of Segregant 1, a glucose-only *FLO8*^{BY} individual, and the *GPA1*^{YJM} Segregant 1 supports *GPA1* as the causal gene underlying the chromosome VIII locus.

suggested that such selective genotyping should have high statistical power (Ehrenreich *et al.* 2010), even in the presence of complex nonadditive genetic effects (Taylor and Ehrenreich 2014), we failed to detect any loci using this strategy (Figure S2A).

FLO8-independent invasion in glucose-only individuals depends on the MAPK cascade

We hypothesized that *FLO8*-independent invasion is genetically heterogeneous in the BYxYJM cross, reducing the statistical power of our genetic mapping effort. To mitigate this potential problem, we attempted to identify causal loci by focusing on different classes of *FLO8*^{BY} segregants. We first looked at *FLO8*^{BY} individuals that showed invasion on both glucose and ethanol, but this analysis did not identify any loci (Figure S2B). We next examined individuals that invaded in only one condition, under the assumption that different mechanisms might underlie condition-specific invasion. Among the segregants showing *FLO8*-independent invasion only on glucose ($n = 36$), nearly all these individuals carried the BY allele of a locus on chromosome VIII, which we were able to delimit to 10 genes (Figure 2A; see also *Materials and Methods*).

To determine the causal gene(s) at the chromosome VIII locus, we replaced the BY allele of each gene in this interval with the YJM allele in a *FLO8*^{BY} segregant that was invasive only on glucose (Segregant 1; see also *Materials and Methods*). Each replacement spanned the promoter, coding region, and part of the downstream region of the tested gene (Figure S1). The only replacement that had an effect was *GPA1*, a subunit of the G-protein-coupled receptor involved in the mitogen-activated protein kinase (MAPK) cascade pheromone response (Fujimura 1989). Converting Segregant 1's *GPA1* allele to the YJM version rendered the strain nearly incapable of invading on glucose and had no effect on ethanol (Figure 2B). BY is known to possess a laboratory-derived amino acid variant (S469I) in *GPA1* that causes a large number of gene expression changes specifically in glucose (Yvert *et al.* 2003; Smith and Kruglyak 2008). This amino acid substitution also may be the causal variant in our study.

Multiple architectures of FLO8-independent invasion in ethanol-only individuals

We next studied *FLO8*^{BY} individuals that were invasive only on ethanol. Because our sample size for this group was small ($n = 12$), we generated backcross populations in a manner similar to Taylor and Ehrenreich (2014) and used these populations to identify loci that influence invasive growth in a single segregant (Segregant 2; see *Materials and Methods*). In the backcross to BY, we screened 192 segregants and found that 16% were invasive only on ethanol. Among these individuals ($n = 30$), we identified a single locus that was nearly fixed for the YJM allele (Figure 3A, top), which was located on chromosome IX and overlapped *FLO11*. *FLO11* is known to harbor extensive functional variation across yeast isolates in both its coding and noncoding regions (Fidalgo *et al.* 2006, 2008). To test for functional variation at *FLO11* in the BYxYJM cross, we separately replaced the coding and noncoding regions of *FLO11* in Segregant 2 with the BY alleles (Figure S1; see also *Materials and Methods*). We found that replacement of the *FLO11* coding region caused a loss of invasion on ethanol (Figure 3B), while replacement of the noncoding region had no effect. A number of amino acid differences, as well as an ~700-bp length difference, distinguish the BY and YJM alleles of *FLO11* (Figure S3 and Figure S4), making it difficult to determine the causal variant.

In the backcross of Segregant 2 to YJM, we also screened 192 segregants and found that 11% were invasive only on ethanol. Among these individuals ($n = 22$), we identified a single locus on chromosome XIV that was fixed for the BY allele. Based on the genotype data, we delimited this interval to 16 candidate genes (Figure 3A, bottom; see also *Materials*

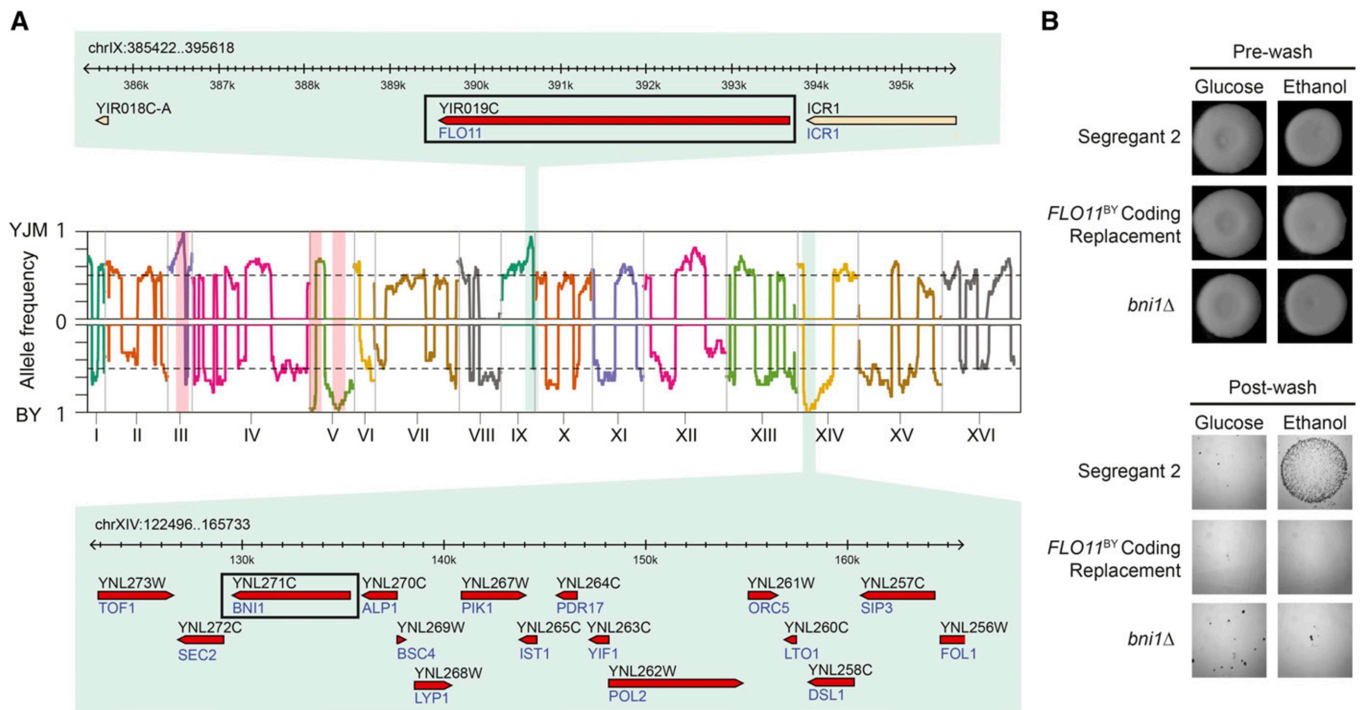

Figure 3 Genetic dissection of ethanol-only invasion by backcrossing Segregant 2 to BY and YJM. (A) Genome-wide relative allele frequency plots for the BY and YJM backcrosses are shown on the top and bottom, respectively. *FLO8* and the markers used to generate haploid progeny are highlighted with red vertical bars, while the strongly enriched intervals on chromosomes IX and XIV are highlighted with green vertical bars. The genomic intervals underlying the chromosome IX and XIV loci are also provided. (B) Comparison of Segregant 2, an ethanol-only *FLO8*^{BY} individual, to *FLO11*^{YJM} replacement and *BNI1* deletion strains in the Segregant 2 background supports *FLO11* and *BNI1* as the causal genes underlying the chromosome IX and XIV loci, respectively.

and Methods). We tested every gene in this interval for an effect on Segregant 2's ability to invade using gene knockouts and found that only deletion of *BNI1* resulted in a loss of invasion (Figure 3B; see also *Materials and Methods*). The BY and YJM alleles of *BNI1* possess 31 coding SNPs, 7 of which are nonsynonymous, as well as 3 SNPs upstream of the gene (Figure S5). Bni1, which has been shown previously to affect invasive growth (Mosch and Fink 1997; Kang and Jiang 2005), is involved in the assembly of actin cables (Sagot *et al.* 2002) and physically interacts with multiple components of the MAPK cascade involved in pheromone response (Chen and Thorner 2007).

Although the *FLO11*^{YJM} coding region contributes to invasion on ethanol, not all the ethanol-only segregants possessed this allele. Among the 12 individuals that were invasive only on ethanol in our genotyped F_2 population, two carried *FLO11*^{BY}. To determine the mechanism that allows these individuals to invade only on ethanol, we backcrossed one relevant segregant (Segregant 3) to BY and YJM. The YJM backcross exhibited very low sporulation; for this reason, we were only able to perform genetic mapping in the BY backcross. We screened 192 segregants and found 32 individuals (17%) that grew invasively only on ethanol. We performed genetic mapping to look for enriched alleles and identified a single locus on chromosome II, at which individuals were fixed for the YJM allele (Figure 4A). This locus was detected at a resolution of four genes, of which only *AMN1* had an effect when deleted. To verify that the BY and YJM alleles functionally differ, we replaced Segregant 3's *AMN1*^{YJM} with *AMN1*^{BY} and found that this resulted in a loss of invasion (Figure 4B and Figure S1; see also *Materials and Methods*). An amino acid variant (D368V) in *AMN1*, which plays a role in daughter cell separation and exit from mitosis (Wang *et al.* 2003), has been implicated as a major determinant of *FLO11*-independent cell clumping in multiple studies (Yvert *et al.* 2003; Li *et al.* 2013) and also may be the causal variant in our study.

Testing for effects of mating type and nongenetic factors on FLO8-independent invasion

Nongenetic factors are known to influence the expression of traits in yeast crosses (*e.g.*, Sirr *et al.* 2015) and also may contribute to *FLO8*-independent invasion. Additionally, because our experiments were conducted exclusively in *MAT***a** haploids, some of the *FLO8*-independent invasion may be mating-type dependent. To test both these possibilities, we generated and sporulated homozygous diploid versions of Segregants 1, 2, and 3 (see *Materials and Methods*). From each individual we obtained 7–10 four-spore tetrads. Only mating type and nongenetic factors should segregate among these spores (see *Materials and Methods*). If we have identified loci that depend on mating type, then invasion should

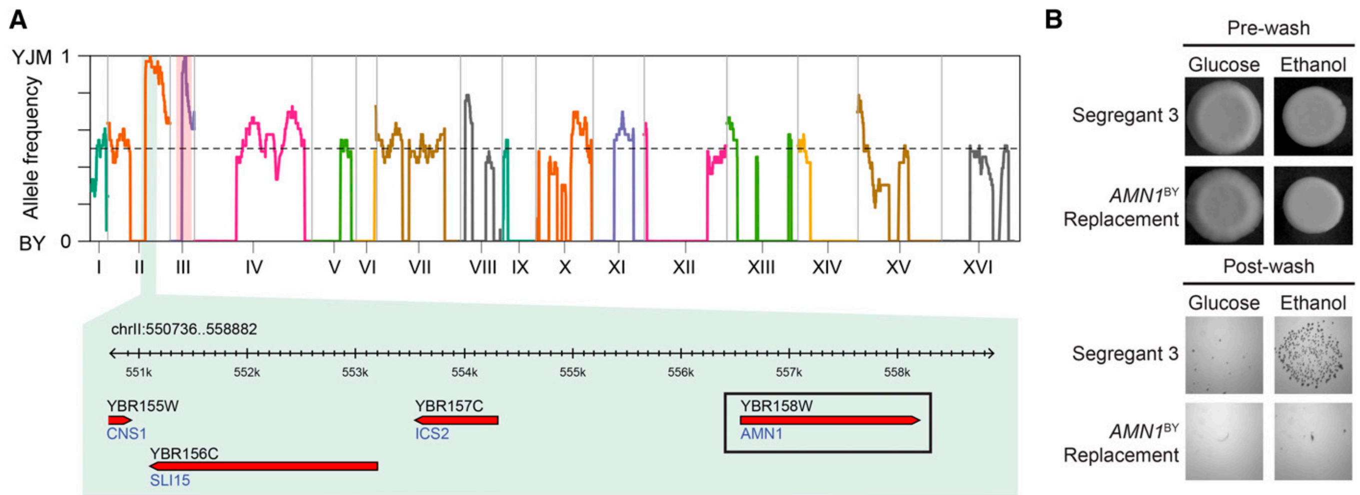

Figure 4 Genetic dissection of *FLO11*-independent ethanol-only invasion by backcrossing of Segregant 3 to BY. (A) Genome-wide relative allele frequency plot of ethanol-only invasion in the backcross of Segregant 3 to BY. The marker used to generate haploid progeny is highlighted with a red vertical bar, while the enriched locus on chromosome II is highlighted with a green vertical bar. The genomic interval underlying the chromosome II locus is also provided. (B) Comparison of Segregant 3, a *FLO11*-independent ethanol-only $FLO8^{BY}$ individual, to $AMN1^{BY}$ replacement strains in the Segregant 3 background supports *AMN1* as the causal gene underlying the chromosome II locus.

cosegregate 2:2 with mating type. Alternatively, if nongenetic factors contribute to *FLO8*-independent invasion, then less than 100% of the examined spores should show the same phenotype as their progenitor.

The effects of mating type and nongenetic factors varied among the tested segregants. For Segregants 2 and 3, which only invade on ethanol, all the haploid spores also showed ethanol-only invasion (Table S7). This indicates that mating type and nongenetic factors likely do not influence the phenotypes of these individuals. In contrast, Segregant 1, which only invades on glucose, provided evidence for both mating-type and nongenetic effects. Among the 40 tested spores from this individual, 16 of 20 *MAT***a** spores showed glucose-only invasion, while none of the 20 *MAT*α spores exhibited invasion (Table S7). This suggests that Segregant 1's phenotype is mating-type dependent and also may have a nongenetic component.

Segregants that invade in a FLO8-independent manner require different transcription factors and cell surface proteins

Our results to this point indicate that *FLO8*-independent invasion has a heterogeneous basis that is largely genetic. This genetic heterogeneity might arise if distinct regulatory factors and/or cell surface proteins facilitate invasion in different segregants and environments. The possibility of such rewiring of invasive growth is supported by recent work showing that the Σ1278b strain requires the transcription factor Tec1 to express *FLO11*, while BY does not (Chin *et al.* 2012), as well as by experiments demonstrating extensive variability in transcription factor binding among progeny from the BYxYJM cross (Zheng *et al.* 2010). Further supporting such a scenario, some of the genes that we cloned have regulatory functions. For example, *GPA1* influences signaling through the MAPK cascade, and the MAPK cascade is known to regulate Ste12, which is a transcriptional activator required for invasion in many pathogenic fungi (Lo and Dranginis 1998; Felden *et al.* 2014).

To explore whether regulatory rewiring might contribute to the genetic heterogeneity in our study, we deleted 11 transcription factors that are known to regulate invasion, as well as *FLO11*, from Segregants 1, 2, and 3 (see *Materials and Methods*). We also performed these deletions in two additional individuals that showed *FLO8*-independent invasion on both glucose and ethanol (hereafter referred to as Segregant 4 and Segregant 5). Although some deletions had quantitative effects on invasion (Figure 5), we focused on cases where deletion of one of the examined genes caused inability to invade. Such complete losses of the phenotype indicate genes that are required for a particular segregant to express *FLO8*-independent invasion.

The examined segregants differed in their requirements for *FLO11* and four transcription factors—*MGA1*, *MSN1*, *RME1*, and *STE12* (Figure 5). None of the deletions caused Segregant 3 to lose its ability to invade, implying that this individual invades in a *FLO11*-independent manner that may not require the examined transcription factors. In contrast, Segregants 1, 2, 4, and 5 showed *FLO11*-dependent invasion but differed in the transcription factors that they require. Segregants 1 and 4 lost the ability to invade when *STE12* was deleted, suggesting that their ability to invade is MAPK dependent. Segregants 2 and 5 required *MSN1*, a transcriptional activator that influences many traits in yeast. While *MSN1* was the only transcription factor that caused loss of invasion in Segregant 2, Segregant 5 also lost its ability to invade when *MGA1* and *RME1* were deleted. The finding that individuals differ in the transcription factors and cell surface

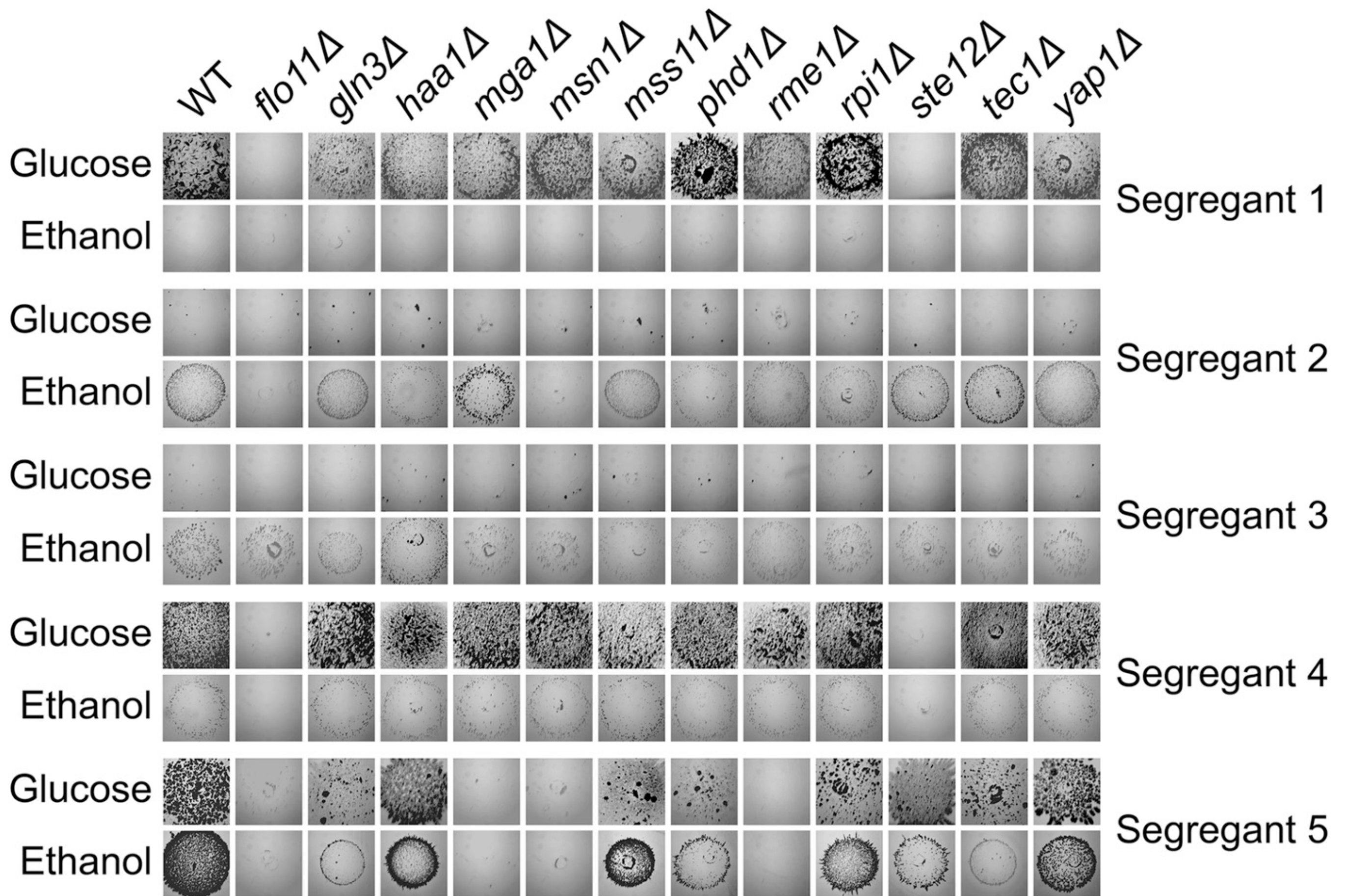

Figure 5 Deletion screen of known *FLO11* activators. *FLO11* and a number of transcription factors that regulate invasive growth were knocked out in Segregants 1–5. These deletion strains then were phenotyped for their ability to invade.

proteins that they require for invasion supports regulatory rewiring as a cause of genetic heterogeneity in our study.

Conclusion

We have shown that a model phenotype in yeast—haploid invasive growth—exhibits extensive nonallelic genetic heterogeneity. This heterogeneity is caused by genetic variants that change the regulation of invasive growth and enable *FLO8*-independent invasion in specific cross progeny. Our results from genetic mapping and genetic engineering experiments suggest that multiple distinct regulatory architectures of *FLO8*-independent invasion segregate in the BYxYJM cross. Although these regulatory architectures require different transcription factors and/or cell surface proteins, they lead to similar abilities to invade.

The present data do not shed light on the specific details of these different regulatory architectures. However, the finding that most BYxYJM segregants that show *FLO8*-independent invasion require *FLO11* suggests that *FLO11* expression is an important component of most of the regulatory architectures. This is of note because *FLO11* has one of the largest promoters in the yeast genome and is thought to be influenced by at least 8 pathways and 15 transcription factors, as well as linked noncoding RNAs and chromatin remodeling complexes (Bruckner and Mosch 2012). The potential of *FLO11* to be regulated by a number of different pathways may facilitate some of the variability in wiring that we have described.

Our finding that different transcription factors and cell surface proteins are required for different genetic backgrounds to invade is similar to the recent discovery of "conditional essential" genes in yeast (Dowell *et al.* 2010). These conditional essential genes are necessary for viability in some isolates but dispensable in others. Our work suggests that conditional essentiality may arise because genetically distinct individuals express similar phenotypes as a result of different underlying regulatory mechanisms. If this is true, then the essentiality of a gene for a trait will depend on which signaling cascade(s) or pathway(s) an individual employs to express a given phenotype in a particular environment.

Given that we have examined a single phenotype in only one pairwise cross and two conditions, we cannot comment on the broader extent of this heterogeneity across species, traits, and environments. However, we note that our results are comparable to recent studies in humans [as summarized in McClellan and King (2010)] and mice (Shao *et al.* 2008;

Spiezio *et al.* 2012), which have shown that many genetic perturbations can produce comparable phenotypic outcomes. To some degree, our effort also represents an integration of previous work describing genetic variation in regulatory pathways (Yvert *et al.* 2003) and transcription factor activity (Zheng *et al.* 2010; Chin *et al.* 2012) across yeast isolates. Importantly, we have extended these past studies by connecting changes in signaling and transcription factor activity, as identified via genetic techniques, to phenotypic outcomes.

Acknowledgments

We thank Jonathan Lee, Martin Mullis, Matthew Taylor, Lars Steinmetz, and two anonymous reviewers for critically reviewing a draft of this manuscript. We also thank Sammi Ali for technical assistance with this project, Oscar Aparicio for the pGAL-*HO* plasmid, Charles Nicolet and the USC Epigenome Center staff and Jinliang Li and the staff at Laragen for their help with Illumina sequencing, and Peter Calabrese for comments on this project during its implementation. Our work was supported in part by grants from the National Institutes of Health (R01GM110255 and R21AI108939), the National Science Foundation (MCB1330874), the Army Research Office (W911NF-14-1-0318), the Alfred P. Sloan Foundation, and the Rose Hills Foundation to I.M.E.

Literature Cited

Bruckner, S., and H. U. Mosch, 2012 Choosing the right lifestyle: adhesion and development in *Saccharomyces cerevisiae*. FEMS Microbiol. Rev. 36: 25–58.

Cao, F., S. Lane, P. Raniga, Z. Zhou, K. Ramon *et al.*, 2006 The Flo8 transcription factor is essential for hyphal development and virulence in *Candida albicans*. Mol. Biol. Cell 17: 295–307.

Chen, R. E., and J. Thorner, 2007 Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochim. Biophys. Acta 1773: 1311–1340.

Chin, B. L., O. Ryan, F. Lewitter, C. Boone, and G. R. Fink, 2012 Genetic variation in *Saccharomyces cerevisiae*: circuit diversification in a signal transduction network. Genetics 192: 1523–1532.

Cullen, P. J., and G. F. Sprague, Jr.., 2000 Glucose depletion causes haploid invasive growth in yeast. Proc. Natl. Acad. Sci. USA 97: 13619–13624.

Dowell, R. D., O. Ryan, A. Jansen, D. Cheung, S. Agarwala *et al.*, 2010 Genotype to phenotype: a complex problem. Science 328: 469.

Ehrenreich, I. M., J. Bloom, N. Torabi, X. Wang, Y. Jia *et al.*, 2012 Genetic architecture of highly complex chemical resistance traits across four yeast strains. PLoS Genet. 8: e1002570.

Ehrenreich, I. M., N. Torabi, Y. Jia, J. Kent, S. Martis *et al.*, 2010 Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature 464: 1039–1042.

Erdeniz, N., U. H. Mortensen, and R. Rothstein, 1997 Cloning-free PCR-based allele replacement methods. Genome Res. 7: 1174–1183.

Felden, J., S. Weisser, S. Bruckner, P. Lenz, and H. U. Mosch, 2014 The transcription factors Tec1 and Ste12 interact with coregulators Msa1 and Msa2 to activate adhesion and multicellular development. Mol. Biol. Cell 34: 2283–2293.

Fidalgo, M., R. R. Barrales, J. I. Ibeas, and J. Jimenez, 2006 Adaptive evolution by mutations in the *FLO11* gene. Proc. Natl. Acad. Sci. USA 103: 11228–11233.

Fidalgo, M., R. R. Barrales, and J. Jimenez, 2008 Coding repeat instability in the *FLO11* gene of *Saccharomyces* yeasts. Yeast 25: 879–889.

Fujimura, H. A., 1989 The yeast G-protein homolog is involved in the mating pheromone signal transduction system. Mol. Cell. Biol. 9: 152–158.

Guo, B., C. A. Styles, Q. Feng, and G. R. Fink, 2000 A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating. Proc. Natl. Acad. Sci. USA 97: 12158–12163.

Halme, A., S. Bumgarner, C. Styles, and G. R. Fink, 2004 Genetic and epigenetic regulation of the *FLO* gene family generates cell-surface variation in yeast. Cell 116: 405–415.

Herskowitz, I., and R. E. Jensen, 1991 Putting the *HO* gene to work: practical uses for mating-type switching. Methods Enzymol. 194: 132–146.

Kang, C. M., and Y. W. Jiang, 2005 Genome-wide survey of non-essential genes required for slowed DNA synthesis-induced filamentous growth in yeast. Yeast 22: 79–90.

Kim, H. Y., S. B. Lee, H. S. Kang, G. T. Oh, and T. Kim, 2014 Two distinct domains of Flo8 activator mediates its role in transcriptional activation and the physical interaction with Mss11. Biochem. Biophys. Res. Commun. 449: 202–207.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Li, J., L. Wang, X. Wu, O. Fang, L. Wang *et al.*, 2013 Polygenic molecular architecture underlying non-sexual cell aggregation in budding yeast. DNA Res. 20: 55–66.

Liti, G., D. M. Carter, A. M. Moses, J. Warringer, L. Parts *et al.*, 2009 Population genomics of domestic and wild yeasts. Nature 458: 337–341.

Liu, H., C. A. Styles, and G. R. Fink, 1996 *Saccharomyces cerevisiae* S288C has a mutation in *FLO8*, a gene required for filamentous growth. Genetics 144: 967–978.

Lo, W. S., and A. M. Dranginis, 1998 The cell surface flocculin Flo11 is required for pseudohyphae formation and invasion by *Saccharomyces cerevisiae*. Mol. Biol. Cell 9: 161–171.

Long, A. D., S. J. Macdonald, and E. G. King, 2014 Dissecting complex traits using the *Drosophila* Synthetic Population Resource. Trends Genet. 30: 488–495.

Manchia, M., J. Cullis, G. Turecki, G. A. Rouleau, R. Uher *et al.*, 2013 The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. PLoS One 8: e76295.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff *et al.*, 2009 Finding the missing heritability of complex diseases. Nature 461: 747–753.

McClellan, J., and M. C. King, 2010 Genetic heterogeneity in human disease. Cell 141: 210–217.

Mosch, H. U., and G. R. Fink, 1997 Dissection of filamentous growth by transposon mutagenesis in *Saccharomyces cerevisiae*. Genetics 145: 671–684.

Nogee, L. M., S. E. Wert, S. A. Proffit, W. M. Hull, and J. A. Whitsett, 2000 Allelic heterogeneity in hereditary surfactant protein B (SP-B) deficiency. Am. J. Respir. Crit. Care Med. 161: 973–981.

Risch, N. J., 2000 Searching for genetic determinants in the new millennium. Nature 405: 847–856.

Rupp, S., E. Summers, H. J. Lo, H. Madhani, and G. Fink, 1999 MAP kinase and cAMP filamentation signaling pathways

converge on the unusually large promoter of the yeast *FLO11* gene. EMBO J. 18: 1257–1269.
Sagot, I., S. K. Klee, and D. Pellman, 2002 Yeast formins regulate cell polarity by controlling the assembly of actin cables. Nat. Cell Biol. 4: 42–50.
Shao, H., L. C. Burrage, D. S. Sinasac, A. E. Hill, S. R. Ernest *et al.*, 2008 Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. Proc. Natl. Acad. Sci. USA 105: 19910–19914.
Sherman, F., 1991 Guide to yeast genetics and molecular, pp. 3–21 in *Methods in Enzymology*, edited by C. Guthrie, and G. R. Fink. Elsevier Academic Press, San Diego.
Sirr, A., G. A. Cromie, E. W. Jeffery, T. L. Gilbert, C. L. Ludlow *et al.*, 2015 Allelic variation, aneuploidy, and nongenetic mechanisms suppress a monogenic trait in yeast. Genetics 199: 247–262.
Smith, E. N., and L. Kruglyak, 2008 Gene-environment interaction in yeast gene expression. PLoS Biol. 6: e83.
Song, Q., C. Johnson, T. E. Wilson, and A. Kumar, 2014 Pooled segregant sequencing reveals genetic determinants of yeast pseudohyphal growth. PLoS Genet. 10: e1004570.
Spiezio, S. H., T. Takada, T. Shiroishi, and J. H. Nadeau, 2012 Genetic divergence and the genetic architecture of complex traits in chromosome substitution strains of mice. BMC Genet. 13: 38.
Storici, F., L. K. Lewis, and M. A. Resnick, 2001 In vivo site-directed mutagenesis using oligonucleotides. Nat. Biotechnol. 19: 773–776.
Sutcliffe, J. S., R. J. Delahanty, H. C. Prasad, J. L. McCauley, Q. Han *et al.*, 2005 Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. Am. J. Hum. Genet. 77: 265–279.
Taylor, M. B., and I. M. Ehrenreich, 2014 Genetic interactions involving five or more genes contribute to a complex trait in yeast. PLoS Genet. 10: e1004324.
Tong, A. H., M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader *et al.*, 2001 Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science 294: 2364–2368.
Wang, Y., T. Shirogane, D. Liu, J. W. Harper, and S. J. Elledge, 2003 Exit from exit: resetting the cell cycle through Amn1 inhibition of G protein signaling. Cell 112: 697–709.
Will, J. L., H. S. Kim, J. Clarke, J. C. Painter, J. C. Fay *et al.*, 2010 Incipient balancing selection through adaptive loss of aquaporins in natural *Saccharomyces cerevisiae* populations. PLoS Genet. 6: e1000893.
Wray, N. R., and R. Maier, 2014 Genetic basis of complex genetic disease: the contribution of disease heterogeneity to missing heritability. Curr. Epidemiol. Rep. 1: 220–227.
Yvert, G., R. B. Brem, J. Whittle, J. M. Akey, E. Foss *et al.*, 2003 *Trans*-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nat. Genet. 35: 57–64.
Zheng, W., H. Zhao, E. Mancera, L. M. Steinmetz, and M. Snyder, 2010 Genetic analysis of variation in transcription factor binding in yeast. Nature 464: 1187–1191.

Communicating editor: L. M. Steinmetz

GENETICS

Supporting Information
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.180661/-/DC1

Regulatory Rewiring in a Cross Causes Extensive Genetic Heterogeneity

Takeshi Matsui, Robert Linder, Joann Phan, Fabian Seidl, and Ian M. Ehrenreich

Copyright © 2015 by the Genetics Society of America
DOI: 10.1534/genetics.115.180661

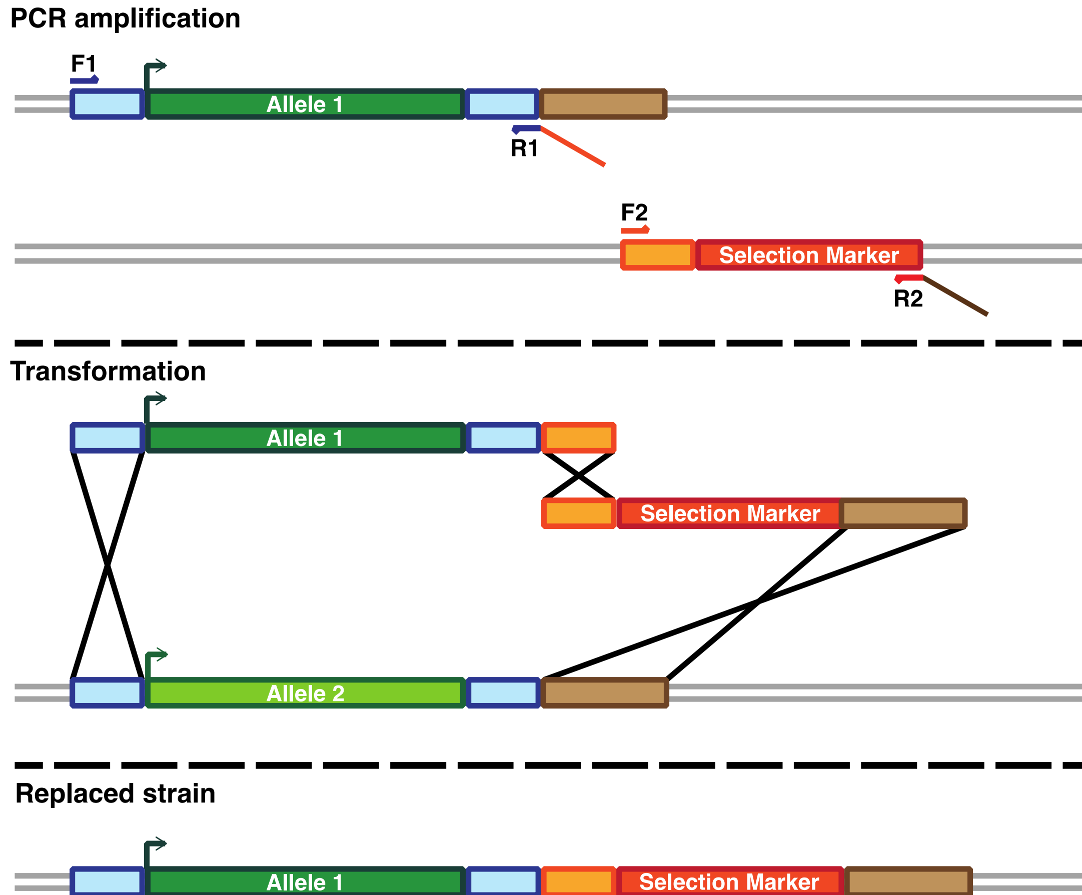

Figure S1. Construction of allele replacements. In the first step, one pair of primers (F1 and R1) was used to amplify the promoter and the coding sequence of the gene to be replaced with 60 bp overlapping the 5' end of the resistance marker attached at the 3' end of the PCR product (shown in orange). Another pair of primers (F2 and R2) was used to amplify the resistance marker with 60 bp overlapping the genomic region immediately downstream of the transcribed potion of the gene using the first primer pair attached at the 3' end of the PCR product. In the second step, the two overlapping PCR products were transformed into the strains. Integration into the genome requires recombination between the PCR products and the target locus.

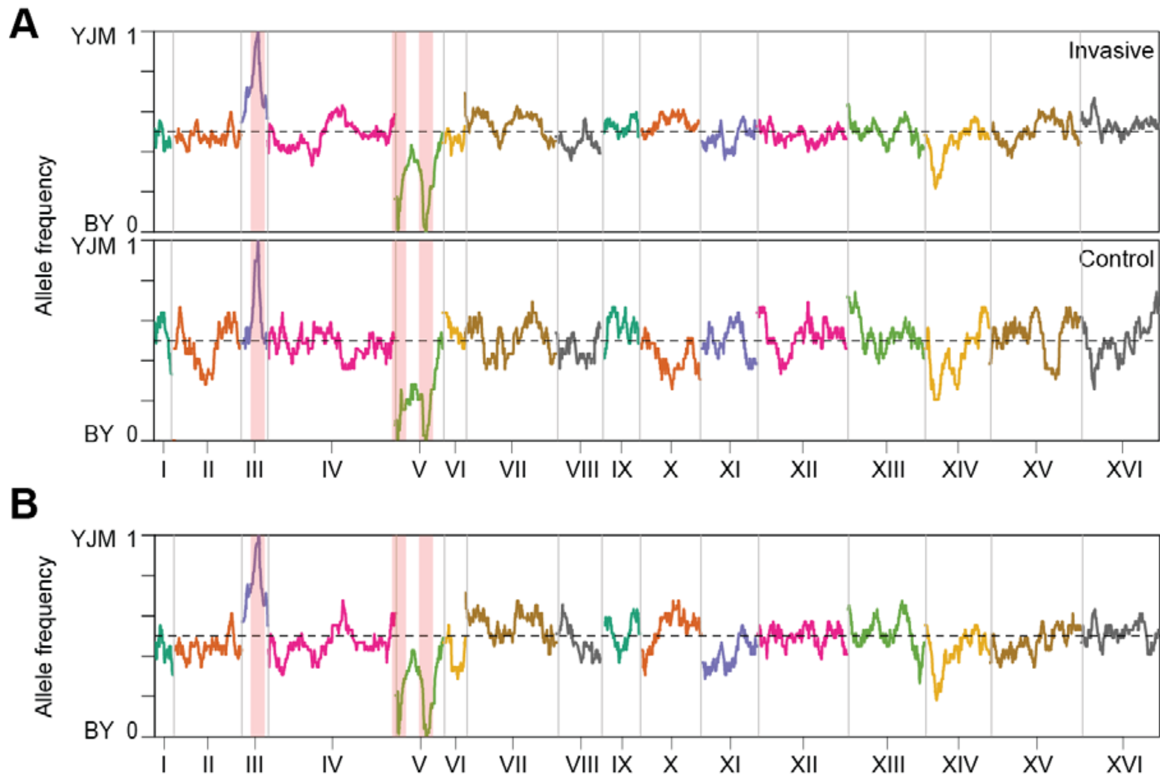

Figure S2. Initial results from selective genotyping of segregants that show *FLO8*-independent invasion. (**A**) Comparison of genome-wide relative allele frequency plot among *FLO8*^{BY} invasive progeny to a non-invasive *FLO8*^{BY} control population. (**B**) Genome-wide relative allele frequency plot among *FLO8*^{BY} segregants that invade on both glucose and ethanol.

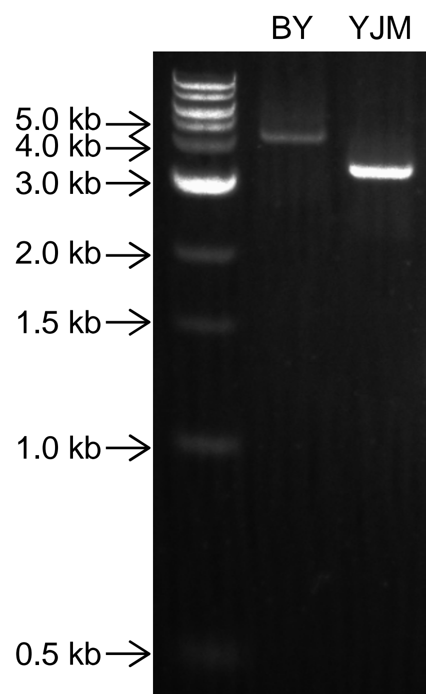

Figure S3. Differences *FLO11* coding region length between BY and YJM. PCR was used to amplify the *FLO11* coding region from the BY and YJM strains. The size of *FLO11*^{BY} was ~4.1kb, while *FLO11*^{YJM} was ~3.4kb.

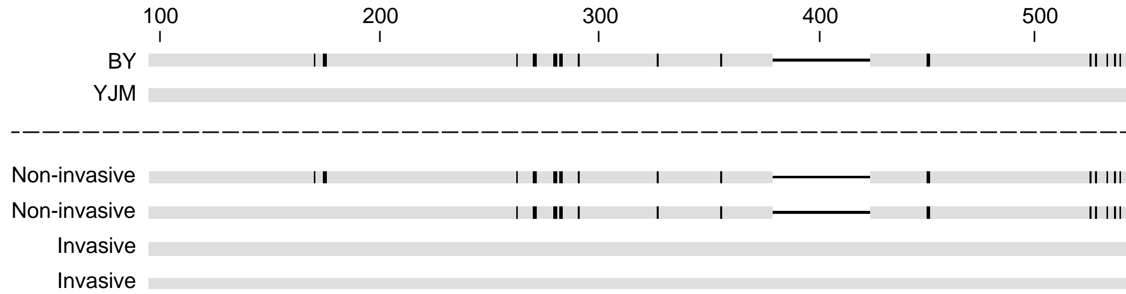

Figure S4. **Replacement of the *FLO11* coding region in segregant 2 with the BY allele causes loss of invasion.** To verify that *FLO11*^{BY} was correctly integrated and replaced using our one-step allele replacement, we PCR amplified the 5' end of the gene, and Sanger sequenced multiple invasive and non-invasive transformants. Only the transformants carrying the BY SNPs (marked in black) toward the 5' end showed loss of invasion, implying that only individuals with most of the *FLO11* gene replaced exhibited loss of invasion. Flo11 protein is comprised of three domains, which are reflected in the sequence of the *FLO11* gene. The N-terminal portion of the protein encodes a hydrophobic signal sequence, is exposed at the cell surface, and binds to ligands. The middle domain largely contains variable length tandem repeats that are enriched for serines and threonines, and is the part of the protein where heavy glycosylation occurs. The C-terminal portion of the protein is a GPI anchor that localizes Flo11 to the cell wall. The highly repetitive nature of the middle portion of *FLO11* makes it difficult to accurately determine the length and sequence of the gene using short Illumina reads. In the regions that we were able to confidently align, we identified 69 SNPs between the BY and the YJM allele, of which 31 were non-synonymous. In addition, we identified that the YJM allele of *FLO11* has a 45bp insertion in the N-terminal region between amino acid position 123 and 124. We also found that no sequencing reads from the YJM mapped to 635 base positions in comparison to BY, which is most likely due to deletions given that the YJM allele of *FLO11* was ~700 bases smaller in comparison to the BY allele (Figure S4). In particular, large stretches of the middle domains were missing from amino acid positions 207 to 315, 359 to 372, 409 to 449, 795 to 808, 824 to 845, and 881 to 899 in the YJM allele. We have not yet determined how these changes alter the functionality of Flo11. We note that this portion of the gene is known to be highly variable across yeast strains, affecting many *FLO11*-dependent traits, such as biofilm formation, flocculation, and invasion.

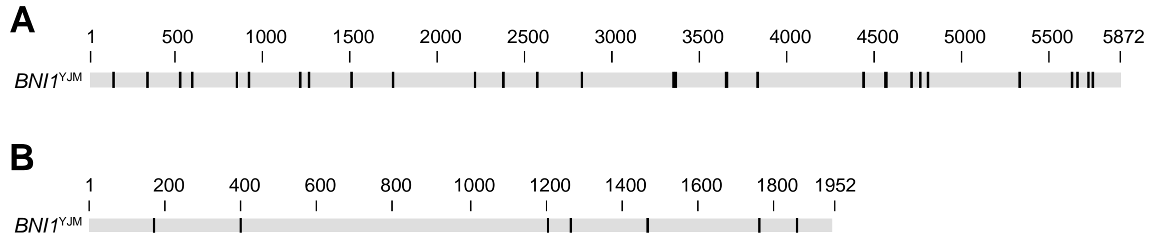

Figure S5. Alignment of the *BNI1* gene. (**A**) Alignment of the nucleotide sequences identified 31 SNPs between the BY and YJM allele of *BNI1*. (**B**) Alignment of the translated amino acid sequence revealed that 7 SNPs were nonsynonymous.

Table S1 Phenotype data and Short Read Archive identifiers for the segregants examined in the paper.

Available for download as an Excel file at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.180661/-/DC1

Tables S2-S6

Available for download as .txt files at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.180661/-/DC1

Table S2. Genotype data for the initial 127 BYxYJM segregants. Genotypes in this table, as well as the following tables, are encoded as 0 for BY and 1 for YJM.

Table S3. Genotype data for the additional 55 BYxYJM segregants that show *FLO8*-independent invasion.

Table S4. Genotype data for the backcross of Segregant 2 to BY.

Table S5. Genotype data for the backcross of Segregant 2 to YJM.

Table S6. Genotype data for the backcross of Segregant 3 to BY.

Table S7. Analysis of dissected tetrads from homozygous diploid derivatives of specific segregants. Phenotypes of spores from homozygous diploid versions of Segregants 1, 2, and 3.

Segregant	Tetrad	MATa spore 1	MATa spore 2	MATalpha spore 1	MATalpha spore 2
1	1	N	N	N	N
1	2	I	N	N	N
1	3	I	N	N	N
1	4	I	I	N	N
1	5	I	I	N	N
1	6	I	I	N	N
1	7	I	I	N	N
1	8	I	I	N	N
1	9	I	I	N	N
1	10	I	I	N	N
2	1	I	I	I	I
2	2	I	I	I	I
2	3	I	I	I	I
2	4	I	I	I	I
2	5	I	I	I	I
2	6	I	I	I	I
2	7	I	I	I	I
3	1	I	I	I	I
3	2	I	I	I	I
3	3	I	I	I	I
3	4	I	I	I	I
3	5	I	I	I	I
3	6	I	I	I	I
3	7	I	I	I	I

I = Invasive, N = Non-invasive