Multiple layers of cryptic genetic variation underlie a yeast complex trait

Jonathan T. Lee*, Alessandro L.V. Coradini*†, Amy Shen*, and Ian M. Ehrenreich*

*Molecular and Computational Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089-2910, USA

†Laboratory of Genomics and Bioenergy, Department of Genetics and Evolution, Institute of Biology, UNICAMP, Campinas, São Paulo, 13083-970, Brazil

Running title: Layers of cryptic genetic variation

Keywords: Cryptic genetic variation, epistasis, genotype-environment interaction, complex traits, genetic architecture

Corresponding author:
Ian M. Ehrenreich
Address: Molecular and Computational Biology Section, Ray R. Irani Hall 201, University of Southern California, Los Angeles, CA 90089-2910, USA
Phone #: 213 – 821 – 5349
Email: ian.ehrenreich@usc.edu

Abstract

Cryptic genetic variation may be an important contributor to heritable traits, but its extent and regulation are not fully understood. Here, we investigate the cryptic genetic variation underlying a *Saccharomyces cerevisiae* colony phenotype that is typically suppressed in a cross of the lab strain BY4716 (BY) and a derivative of the clinical isolate 322134S (3S). To do this, we comprehensively map the trait's genetic basis in the BYx3S cross in the presence of three different genetic perturbations that enable its expression. This allows us to detect and compare the specific loci harboring cryptic genetic variants that interact with each perturbation. In total, we identify 21 loci, all but one of which interacts with just a subset of the perturbations. Beyond impacting which loci contribute to the trait, the genetic perturbations also influence the extent of additivity, epistasis, and genotype-environment interaction among the detected loci. Additionally, we show that the single locus interacting with all three perturbations corresponds to the coding region of the cell surface gene *FLO11*. Nearly all of the other loci influence *FLO11* transcription in *cis* or *trans*. However, the perturbations reveal cryptic genetic variation in different pathways and sub-pathways upstream of *FLO11*, suggesting that multiple layers of cryptic genetic variation with highly contextual effects underlie the trait. Our work demonstrates an abundance of cryptic genetic variation in transcriptional regulation and illustrates how this cryptic genetic variation complicates efforts to study the relationship between genotype and phenotype.

Article Summary

To better understand cryptic genetic variation, we comprehensively map the genetic basis of a trait that is suppressed in a yeast cross. By examining how three different genetic perturbations give rise to the trait, we identify 21 loci harboring cryptic genetic variants, nearly all of which play roles in the transcriptional regulation of the same key gene. We examine these loci in greater detail and show that the perturbations affect every aspect of the trait's genetic architecture, including which cryptic genetic variants have phenotypic effects, as well as the extent of additivity, epistasis, and genotype-environment interaction among these variants.

Introduction

Most research on complex traits focuses on characterizing the genetic basis of phenotypic diversity that is visible within populations (ATWELL *et al.* 2010; AYLOR *et al.* 2011; MACKAY *et al.* 2012; BLOOM *et al.* 2013). Yet, these same populations can also harbor cryptic genetic variation that does not typically impact phenotype, and is only observable when particular genetic or environmental perturbations occur (RUTHERFORD AND LINDQUIST 1998; QUEITSCH *et al.* 2002; BERGMAN AND SIEGAL 2003; DWORKIN *et al.* 2003; GIBSON AND DWORKIN 2004; JAROSZ AND LINDQUIST 2010; GEILER-SAMEROTTE *et al.* 2016). This cryptic variation may be an important source of phenotypic variability in medically and evolutionarily significant traits (LE ROUZIC AND CARLBORG 2008; MCGUIGAN AND SGRO 2009; PAABY AND ROCKMAN 2014). Thus, it is imperative that we determine the extent of cryptic variation within populations, as well as the mechanisms that convert this cryptic variation between silent and visible states. However, such work is inherently difficult because the specific perturbations needed to uncover cryptic variation, as well as the exact identities of the cryptic genetic variants (hereafter, 'cryptic variants') that are affected by these perturbations, are rarely known. Such information is critical to obtaining a more complete, mechanistic understanding of cryptic variation.

In previous papers, we developed an experimental system that can be used to systematically identify cryptic variants influencing a colony phenotype in *Saccharomyces cerevisiae* (TAYLOR AND EHRENREICH 2014; TAYLOR AND EHRENREICH 2015b; LEE *et al.* 2016; TAYLOR *et al.* 2016). The lab strain BY4716 (BY), a haploid derivative of the clinical isolate 322134S (3S), and their haploid recombinant progeny form 'smooth' colonies when grown on solid media (**Figure 1A**). However, certain *de novo* or induced mutations, as well as recombination between the promoter and coding region of the cell surface gene *FLO11*, can enable some BYx3S segregants to express an alternative 'rough' colony phenotype (**Figure 1B**). Throughout the current paper, we refer to these new alleles that are not naturally present in BY or 3S, but make it possible for the rough phenotype to be expressed in the BYx3S cross as genetic perturbations (or 'GPs'). These GPs on their own are insufficient to cause expression of the rough phenotype; rather, cryptic variants that segregate between BY and 3S are also needed. Moreover, BYx3S segregants that express the trait in the presence of a given GP usually exhibit the phenotype in a temperature-sensitive manner. However, by backcrossing these segregants and examining their backcross progeny

59 across temperatures, additional cryptic variation can often be found that reduces or eliminates
60 temperature sensitivity (**Figure 1C**). In other words, expression of the trait results from complex
61 interactions between the GPs, segregating loci, and temperature.
62
63 Crucially, genetic mapping and genetic engineering techniques can be used to comprehensively
64 identify the specific GPs and cryptic variants that enable a given rough segregant to express the
65 trait. In our initial work on colony morphology in the BYx3S cross, we found that a *de novo* loss-
66 of-function mutation in *IRA2* ('GPa'), a negative regulator of Ras signaling, had occurred during
67 the generation of BYx3S segregants and enabled ~2% of cross progeny to express the phenotype
68 (TAYLOR AND EHRENREICH 2014) (**Figure 1D**). Across several papers, we demonstrated that GPa
69 causes trait expression through temperature-dependent, higher-order genetic interactions
70 involving cryptic variants inherited from both BY and 3S (TAYLOR AND EHRENREICH 2014;
71 TAYLOR AND EHRENREICH 2015b; LEE *et al.* 2016). Subsequently, we identified other *de novo*
72 and induced mutations that also facilitate expression of the rough phenotype in the BYx3S cross
73 (TAYLOR AND EHRENREICH 2015b; TAYLOR *et al.* 2016). These other mutations tended to also
74 disrupt negative regulation of signaling and transcription within the Ras pathway, and mostly
75 interacted with the same alleles found in the studies focused on GPa.
76
77 In a screen of 106 independently generated BYx3S crosses in which 17% of the crosses
78 produced at least one rough segregant, we also found individuals that expressed the trait despite
79 lacking any detectable *de novo* mutations (TAYLOR *et al.* 2016). Instead, two-thirds of these
80 segregants inherited recombination events in the promoter of *FLO11* ('GPb', **Figure 1E**). In
81 total, we recovered six different recombination breakpoints, which all occurred within 1.3 kb of
82 each other. *FLO11* encodes a flocculin whose display on the cell surface facilitates cell-cell
83 adhesion; this protein is required for expression of the rough phenotype (LO AND DRANGINIS
84 1996; TAYLOR AND EHRENREICH 2015b). Genetic engineering experiments showed that the
85 recombination events brought at least two alleles at closely linked loci in or near *FLO11* onto the
86 same chromosome, resulting in a new *FLO11* haplotype that behaves like the mutations
87 described in our earlier studies (TAYLOR *et al.* 2016). Specifically, segregants with a BY
88 promoter and a 3S coding region had the potential to express the trait. Despite discovering GPb
89 in this past study, we neither resolved the cryptic variant(s) in the *FLO11* promoter nor

90 comprehensively mapped the loci enabling this GP to exert a phenotypic effect. Thus, it was not
91 possible to compare the genetic basis of the rough phenotype in the presence of GPb to our initial
92 work on GPa.
93
94 Here, we use the rough colony system to determine how different types of GPs interact with
95 distinct cryptic variants in the BYx3S cross to produce the same trait. In addition to GPa and
96 GPb, we also examine a third GP that facilitates expression of the rough phenotype. To generate
97 'GPc,' we knocked out the activator Flo8 and the repressor Sfl1 (**Figure 1G**), which are the main
98 transcription factors that act downstream of the Ras pathway to regulate colony morphology in
99 the cross. Previously, we showed that deletion of *SFL1* is sufficient to enable Ras-dependent
100 cryptic variants to express *FLO11* and that transcription of *FLO11* in these *sfl1*Δ segregants is
101 Flo8-dependent (TAYLOR AND EHRENREICH 2015b). By eliminating these key regulators and
102 screening for segregants expressing the trait, we sought to uncover previously unidentified
103 cryptic variants that can also give rise to the rough phenotype. We successfully recovered one
104 rough BYx3S *flo8*Δ *sfl1*Δ segregant, making it possible to examine the genetic basis of the
105 phenotype in the absence of its primary transcriptional regulators.
106
107 In this paper, we comprehensively determine the genetic basis of the rough phenotype across
108 temperatures for GPb and GPc, and compare these results to our past work on GPa. Across the
109 three GPs, we identify 21 loci that contribute to the rough phenotype. Of these loci, 20 show
110 conditional phenotypic effects that are influenced by particular combinations of GP, other loci,
111 and temperature. Although all three factors prove important, GP is by far the strongest
112 determinant of which loci show phenotypic effects, impacting nearly all identified loci. In fact,
113 12 of the detected loci only exert phenotypic effects in the presence of a single GP. Additionally,
114 we find that the identified loci exhibit varying degrees of additivity, epistasis, and genotype-
115 environment interaction depending on the GP that uncovers them. At the molecular level, most,
116 if not all, of the identified loci influence *FLO11* regulation in *cis* or *trans*, suggesting that our
117 findings result from complex genetic and environmental effects on the regulation of a single key
118 gene, *FLO11*. These findings enhance our understanding of both the extent of cryptic variation
119 within populations and the mechanisms by which GPs reveal cryptic variants. Further, they show
120 how conditional uncovering of cryptic genetic variation can result in highly divergent genetic

121 architectures that produce the same trait, even within a single population examined in a common
122 environment.
123
124 **Materials and Methods**
125 *Knockout of FLO8* and *SFL1 in BY and 3S*
126 *FLO8* and *SFL1* targeting guide RNA (gRNA) sequences were cloned into the pML104 plasmid
127 vector, which carries Cas9 and *URA3* (LAUGHERY *et al.* 2015). Each resulting plasmid was
128 transformed into BY or 3S alongside a double-stranded oligo repair template for homologous
129 recombination. The *flo8 sfl1*Δ BY and 3S strains were both generated through a series of two
130 sequential knockouts. The lithium acetate method (GIETZ AND WOODS 2002) was used for
131 transformation. Oligos were 90 bases long with homology to either *FLO8* or *SFL1*, with
132 inclusion of a stop codon followed by single base frameshift deletion in the middle. Stop codons
133 were introduced at amino acids 155 and 39 in *FLO8* and *SFL1*, respectively. Transformed cells
134 were plated on solid yeast nitrogen base (YNB) media lacking uracil to select for retention of the
135 pML104 plasmid. Plasmids were then eliminated from the transformants by plating on 5-
136 fluoroorotic acid (5-FOA). Presence of the intended genetic modifications was checked in the
137 transformants using Sanger sequencing.
138
139 *Isolation of rough BYx3S F_2 segregants*
140 GPa and GPb were identified in BYx3S crosses produced by sporulating independently
141 generated BY/3S diploids. The GPa rough F_2 segregant included in the current study is discussed
142 in (TAYLOR AND EHRENREICH 2014; TAYLOR AND EHRENREICH 2015b; LEE *et al.* 2016), while
143 the GPb rough F_2 segregant used in this study was described in (TAYLOR *et al.* 2016), in which it
144 was referred to as 'Rough segregant 13'. GPa and GPb were identified in these past studies by
145 taking individual rough F_2 segregants recovered from screens of BYx3S crosses, performing
146 genetic mapping in backcross populations obtained by mating the respective rough F_2 segregants
147 to both BY and 3S, and analyzing whole genome sequencing data for backcross segregants. To
148 obtain a rough F_2 segregant with GPc, *flo8*Δ *sfl1*Δ BY and 3S strains were mated to one another.
149 The resulting diploid was sporulated and plated at low density (~300 colonies per plate) onto
150 YNB agar containing canavanine to select for MAT**a** haploids using the synthetic genetic array
151 system (TONG *et al.* 2001). After five days at 30°C, colonies were replicated onto yeast extract-

152 peptone-ethanol (YPE) plates. A single rough colony was identified after four days of growth at
153 21°C.
154
155 *Generation of backcross segregants*
156 All data for GPa was described in previous papers (TAYLOR AND EHRENREICH 2014; TAYLOR
157 AND EHRENREICH 2015b; LEE *et al.* 2016), while data for GPb and GPc were generated in this
158 current paper. To obtain backcross progeny for genetic mapping and segregation analysis, the
159 rough F_2 segregant with GPb was backcrossed to wild type BY and 3S strains, while the rough F_2
160 segregant with GPc was backcrossed to *flo8*Δ *sfl1*Δ BY and 3S strains. A second-generation
161 backcross was also performed for GPb; specifically, a rough segregant from the GPbx3S cross
162 was mated to BY. For all backcrosses, diploids were sporulated and plated at low density onto
163 YNB plates containing canavanine to select for random MAT**a** spores using the Synthetic
164 Genetic Array marker system (TONG *et al.*). After five days of growth at 30°C, haploid colonies
165 were replicated onto YPE plates and incubated at 21°C. Backcross segregants expressing rough
166 morphology after four days of growth on YPE were inoculated into liquid yeast extract-peptone-
167 dextrose (YPD) media and grown overnight at 30°C. Freezer stocks of rough segregants were
168 made by mixing aliquots of these cultures with 20% glycerol solution and storing these glycerol
169 stocks at -80°C.
170
171 *Phenotyping at multiple temperatures*
172 Cells from freezer stocks were inoculated into liquid YPD media, grown for two days at 30°C,
173 and then pinned onto three YPE plates. Each plate was then incubated at a single, constant
174 temperature: 21, 30, or 37°C. The colonies were then phenotyped for the colony trait after four
175 days of growth, and designated as belonging to one of three temperature sensitivity classes:
176 expression of the trait at 21°C only, expression of the trait at 21 and 30°C only, or expression of
177 the trait at all examined temperatures. Genetic mapping was then performed separately on each
178 of these temperature sensitivity classes.
179
180 *Genotyping of GPb and GPc rough segregants*
181 Segregants from each temperature sensitivity class were inoculated into liquid YPD. DNA was
182 extracted from overnight cultures using the Qiagen DNeasy 96 Blood and Tissue kit. Illumina

183 sequencing libraries were then prepared using the Illumina Nextera kit, with a unique pair of
184 dual-indexed barcodes for each individual. Between 48 and 156 segregants from each
185 combination of GP, backcross, and temperatures sensitivity were sequenced. Sequencing was
186 performed at the Beijing Genomics Institute on an Illumina HiSeq 4000 using paired-end 100 bp
187 x100 bp reads. Each segregant was sequenced to an average per site coverage of at least 2.5x.
188 Reads were aligned to either a BY or 3S reference genome using the Burrows-Wheeler Aligner
189 (BWA) version 7 with options mem -t 20 (LI AND DURBIN 2009). SAMtools was used to
190 generate mpileup files (LI *et al.* 2009), and genome-wide allele frequencies were calculated at
191 36,756 SNPs that were previously identified between BY and 3S (TAYLOR *et al.* 2016).
192
193 *Genetic mapping of loci underlying the rough phenotype*
194 Segregants' genotypes at each SNP were determined using a Hidden Markov model,
195 implemented in the HMM package in R. Loci associated with the trait for each combination of
196 GP, backcross, and temperature sensitivity were identified using binomial tests. Sites were
197 considered statistically significant at a Bonferroni-corrected threshold of $p < .01$. Multiple-
198 testing correction was performed on each backcross population on its own, as the number of
199 unique tests varied 837 to 1,526 among these populations. We defined the interval surrounding
200 identified loci by computing the -log10(*p*-value) at each linked SNP and determining the SNPs at
201 which this statistic was 2 lower than the peak marker at a given locus. These bounds were used
202 in fine-mapping and comparison of loci detected in different combinations of GP, backcross, and
203 temperature sensitivity.
204
205 *Testing for genotypic heterogeneity*
206 Observed and expected haplotype frequencies for each pair of SNPs were compared using χ^2
207 tests in a custom Python script. Expected haplotype frequencies were calculated as the product of
208 the individual allele frequencies at each of the two sites. To reduce the number of statistical tests,
209 SNPs containing the same information across all segregants in a given backcross population
210 were collapsed into a single marker. The Benjamini-Hochberg method for False Discovery Rate
211 (FDR) (BENJAMINI AND HOCHBERG 1995) was then implemented using the statsmodels Python
212 module (SEABOLD AND PERKTOLD 2010). A stringent FDR of 0.0001 was employed and regions

213 of the genome within 30,000 bases of the ends of chromosomes were excluded to avoid
214 genotyping errors.
215
216 *Exploration of additivity and epistasis*
217 Expected multi-locus genotype frequencies were calculated as the product of the frequencies of
218 the individual alleles segregating in a given backcross. For each population, observed and
219 expected genotype frequencies were compared using a χ^2 test with degrees of freedom equal to
220 one less than the number of possible genotypes. These analyses only included loci where both
221 the BY and 3S allele were present among the rough segregants obtained from a population, and
222 were only performed on backcross populations in which more than one multi-locus genotype was
223 present.
224
225 *Genetic engineering at the FLO11 promoter and other loci*
226 Gene deletions were generated using the *kanMX* cassette (WACH *et al.* 1994) and lithium acetate
227 transformation (GIETZ AND WOODS 2002). A genomic region of interest was replaced with a
228 PCR amplicon of the *kanMX* cassette that, on each end, had 60 bases of homology to the targeted
229 gene. Transformed cells were plated on YPD + G418 to select for integration of *kanMX* and
230 insertion was verified using PCR.
231
232 Allele replacements were performed using a two-step, CRISPR/Cas9-mediated approach. A
233 *kanMX* deletion strain, generated as described above, was transformed with the pML104 plasmid
234 (LAUGHERY *et al.* 2015) carrying the Cas9 gene and a gRNA sequence targeting *kanMX*, along
235 with a PCR product repair template that should result in replacement of *kanMX* with the desired
236 allele. Cells were plated on YNB plates lacking uracil to select for retention of the plasmid and
237 then plated again onto 5-FOA to eliminate the plasmid. Replacement of the *kanMX* cassette was
238 verified by PCR and Sanger sequencing, and transformants were plated and screened on YPE to
239 determine the phenotypic effects of allele replacement. Also, in parallel with each allele
240 replacement, we generated control strains where *kanMX* was replaced with the original sequence
241 at a given site. The phenotypes of allele replacement strains were then compared to the
242 phenotypes of these control strains that were generated in parallel. Engineering of the *FLO11*
243 promoter was performed in the BYx3S F_2 segregant with GPb. Gene deletions and allele

244 replacements for other genes were performed in a representative BY or 3S backcross segregant
245 harboring either GPb or GPc.
246
247 **Results**
248 *Isolation of a rough flo8Δ sfl1Δ segregant*
249 Our past work showed that expression of the rough phenotype in the BYx3S cross occurs due to
250 genetic interactions between GPs that were not present in the cross parents and cryptic variants
251 that segregate in the cross and mainly reside in the Ras pathway (TAYLOR AND EHRENREICH
252 2014; TAYLOR AND EHRENREICH 2015b; LEE *et al.* 2016; TAYLOR *et al.* 2016). To examine
253 whether genetic variation beyond the Ras pathway might also be able to contribute to the trait,
254 we knocked out *FLO8* and *SFL1* in both BY and 3S using CRISPR/Cas9 (**Methods**). We then
255 employed random spore techniques to generate and screen > 100,000 *flo8*Δ *sfl1*Δ (GPc) cross
256 progeny for the trait at 21°C (**Methods**). We used this condition because the trait is less
257 genetically complex at 21°C than at higher temperatures, making it easier to find rough
258 segregants in a screen (LEE *et al.* 2016). The GPc screen produced a single rough segregant,
259 implying that expression of the rough phenotype in the absence of Flo8 and Sfl1 requires
260 multiple alleles that segregate in the cross, some of which must be inherited from BY and others
261 of which must be inherited from 3S.
262
263 *The three GPs vary in their potential to express the phenotype across temperatures*
264 In our previous work, we showed that GPa segregants typically express the trait in a temperature
265 sensitive manner, but that cryptic variation in the cross can eliminate this temperature sensitivity
266 for some individuals (LEE *et al.* 2016). Because our GPb and GPc rough segregants were both
267 obtained at 21°C, we assessed whether GPb and GPc strains also have the potential to express
268 the trait at higher temperatures. To check this, we backcrossed GPb and GPc F_2 segregants to
269 both BY and 3S, and then phenotyped the resulting progeny at 21, 30, and 37°C (**Figure 1H**).
270 Note, such backcrossing generates new multi-locus genotypes that may have the potential to
271 express the rough phenotype at higher temperatures than the F_2 segregants recovered from initial
272 screens.
273

Upon examining 768 GPb and GPc backcross progeny at higher temperatures, we found that some GPb backcross segregants expressed the trait at higher temperatures whereas GPc backcross segregants did not (**Figure S1**). We observed varying degrees of temperature sensitivity among the rough GPb segregants, with only a minority of rough individuals capable of expressing the trait at all examined temperatures (**Figure S1**). In our past work on GPa (LEE *et al.* 2016), we also found that robustness to temperature was less frequent than temperature sensitive trait expression. This was because ability to robustly express the rough phenotype across temperatures involved more loci than temperature sensitive trait expression. Thus, when considered in light of our past findings, our current results suggest that, in the presence of GPb, ability to express the rough phenotype across temperatures is also more genetically complex than ability to express the trait only at lower temperatures. These findings also show that despite making it possible for the rough phenotype to be expressed at 21°C, GPc provides an inherently limited potential for trait expression at higher temperatures. This may be because Flo8 and Sfl1 are necessary for expression of the rough phenotype at 30 and 37°C.

The GPs uncover distinct loci

To map loci involved in expression of the rough phenotype across temperatures, we generated > 60,000 and > 12,000 GPb and GPc backcross segregants, respectively. First-generation backcross segregants were used exclusively in mapping, except in the case of the 3S backcross of GPb at high temperature. Because expression of the rough phenotype at 30 and 37°C is rare among GPbx3S backcross segregants (**Figure S1**), determining the genetic basis of the trait among GPbx3S backcross segregants required combining information from the first-generation backcross and a second-generation backcross. Specifically, a rough first-generation GPbx3S backcross segregant was mated to 3S (**Figure S2**), which increased the frequency of the trait among backcross progeny by halving the proportion of the genome that was segregating. GPb segregants were phenotyped at 21, 30 and 37°C, while GPc segregants were only phenotyped at 21°C. Low-coverage whole genome sequencing was performed on between 48 and 151 individuals for each GP–backcross–temperature combination, and loci associated with the trait were identified based on their enrichment among genotyped segregants (**Methods**).

304 In past work (Figure 5 in (LEE *et al.*)), we identified eight loci that act in different multi-locus
305 genotypes to enable GPa segregants to express the trait at particular temperatures (**Figure 2**). BY
306 and 3S contributed the causal alleles at two and four of these loci, respectively, while the
307 remaining two loci were detected in both the BY and 3S allele states in different multi-locus
308 genotype and temperature contexts (**Figure 3**). Here, genetic mapping focused on GPb
309 segregants detected a total of 11 loci across the three environments (**Figure 2**, **Figure S3**, **Table**
310 **S1**). Among these loci, six and three were detected in the BY and 3S allele states, respectively
311 (**Table 1**). The other two loci were identified in both the BY and 3S allele states, depending on
312 the multi-locus genotype context in which they occurred. Eight of these loci influenced the trait's
313 expression independent of temperature. Of the remaining loci, two were detected among
314 individuals that could express the trait at up to 30°C, while the third was identified among
315 individuals that could express the trait at up to 37°C. In addition, we identified 12 loci in the
316 presence of GPc (**Figure 2**, **Figure S3**, **Table S1**). Among the loci found in the presence of GPc,
317 seven were contributed by BY and five were contributed by 3S (**Table 1**).
318
319 Some of the loci detected across GPs and temperatures both overlapped and showed involvement
320 of the same parental allele, indicating potential involvement of the same underlying cryptic
321 variants. Supporting this possibility, each locus also corresponded to a causal gene that we had
322 cloned during our work on GPa (TAYLOR AND EHRENREICH ; TAYLOR AND EHRENREICH ; LEE *et*
323 *al.* ; TAYLOR *et al.*). Assuming that these overlapping regions represent the same cryptic variant,
324 we consolidated all detected genomic intervals into 21 distinct loci (**Figure 2**, **Table 1**). Among
325 these loci, one was found in the presence of all three GPs, 8 were found in the presence of two
326 GPs, and 12 were found in the presence of a single GP. This indicates that nearly all of the
327 detected loci show differential responsiveness to the GPs and that the majority of the loci we
328 have identified are cryptic variants that only act in the presence of specific GPs. However, if our
329 assumption of overlapping loci representing the same cryptic variants is incorrect, then the
330 responsiveness of the identified cryptic variants to GP and temperature would be even higher
331 than discussed in this paper.
332
333 *The GPs alter genotype-environment interaction*

334 We assessed how identified loci interact with particular GPs, each other, and temperature to
335 produce the rough phenotype. Given that we performed such an analysis on GPa in the past (LEE
336 *et al.* 2016) and GPc segregants can only express the trait at 21°C, we focused this analysis on
337 GPb. Note, our past work on GPa found that at particular levels of temperature sensitivity,
338 distinct sets of epistatic cryptic variants form multi-locus genotypes that produce the phenotype
339 (LEE *et al.* 2016). For example, at 30°C, two distinct combinations of four and five cryptic
340 variants act in conjunction with GPa to cause the trait's expression (**Figure 3**). Examination of
341 first-generation GPb backcross segregants produced results comparable to our findings for GPa.
342 Among GPb segregants expressing the trait exclusively at 21°C, we detected a pair of interacting
343 loci on chromosomes XIII and XV, which then allowed us to identify distinct multi-locus
344 genotypes present among these individuals (**Figure 3**, **Figure S4**, **Note S1**, **Methods**). Similar to
345 our work on GPa (LEE *et al.* 2016), only one of the GPb multi-locus genotypes found at 21°C
346 provided a foundation upon which additional allele substitutions can facilitate trait expression at
347 higher temperatures (**Figure 3**, black lines). Among these GPb segregants exhibiting the trait at
348 higher temperatures, single multi-locus genotypes indicative of higher-order epistasis facilitated
349 trait expression at temperatures up to 30 and 37°C (**Figure 3**). Comparison our results for GPb to
350 our past work on GPa found that more than half of the loci and the exact multi-locus genotypes
351 differed between the two GPs (**Figure 3**). These results show that the GPs significantly modify
352 genotype-environment interactions underlying the rough phenotype.
353
354 *The GPs affect additivity and epistasis among identified loci*
355 The only temperature at which all three GPs enable trait expression is 21°C. To compare the
356 quantitative genetic architectures enabling the three GPs to induce the rough phenotype at this
357 temperature, we used genotype data from backcrosses (**Figure 4A** through **C**, **Methods**). If loci
358 act in a predominantly additive manner in the presence of a given GP, observed multi-locus
359 genotype frequencies should match expected multi-locus genotype frequencies, which can be
360 calculated as the product of the frequencies of every individual allele involved in a multi-locus
361 genotype. In contrast, if epistasis meaningfully contributes to the trait, observed multi-locus
362 genotype frequencies should depart from expected multi-locus genotype frequencies. Based on
363 these tests, we observed significant deviation from expected frequencies for GPa (**Figure 4A**)
364 and GPb (**Figure 4B**), confirming that epistasis plays a significant role in the trait in the presence

365 of these GPs. However, data for GPc suggested that trait expression in the presence of this GP is
366 entirely additive (**Figure 4C**). These findings imply that although the three GPs in this study can
367 each enable expression of the same rough phenotype, they do so not only through largely distinct
368 cryptic variants but also through fundamentally different quantitative genetic architectures.
369
370 *Coding and regulatory variation in FLO11 plays an essential role in the trait*
371 Only a single locus exhibited a phenotypic effect in the presence of all three GPs. This locus
372 corresponds to the 3S allele of the *FLO11* coding region (LEE *et al.* 2016). The BY and 3S
373 alleles of *FLO11* possess not only 57 synonymous and 29 nonsynonymous SNP differences, but
374 also a length polymorphism of ~581 nucleotides (**Figure S5)**. $FLO11^{3S}$ encodes a longer cell-
375 surface protein than $FLO11^{BY}$, an attribute that is known to favor the expression of multi-cellular
376 traits like the rough phenotype (VERSTREPEN *et al.* 2005; FIDALGO *et al.* 2006; FIDALGO *et al.*
377 2008; ZARA *et al.* 2009; HOPE AND DUNHAM 2014; MATSUI *et al.* 2015). Although this length
378 polymorphism is likely causal for the trait, we cannot rule out the possibility that some of the
379 SNPs also play a role.
380
381 In addition to being a required component of GPb, we found the BY allele of the *FLO11*
382 promoter is necessary for GPc segregants to express the trait (**Figure 3**, **Table 1**). *FLO11* has
383 one of the largest and most complex promoters in *S. cerevisiae*, with more than 17 transcription
384 factors and 6 signaling cascades capable of influencing its regulation (BRUCKNER AND MOSCH
385 2012). Through genetic engineering experiments, we localized the causal variant in the *FLO11*
386 promoter to a Rim101 binding site that is present in 3S but not BY (**Figure 5)**. This finding is
387 consistent with the important role that transcriptional derepression of *FLO11* plays in expression
388 of the rough phenotype (TAYLOR AND EHRENREICH 2015b; TAYLOR *et al.* 2016). We note that
389 although Rim101 has been described as a *FLO11* activator in other strains of *S. cerevisiae*, this
390 role is indirect and mediated through its role in silencing *NRG1* (KUCHIN *et al.* 2002; LAMB AND
391 MITCHELL 2003; BARRALES *et al.* 2008), which encodes a repressor that directly binds the
392 *FLO11* promoter. The Rim101 binding site in the 3S *FLO11* promoter most likely results in
393 direct repression of *FLO11* by Rim101, which reinforces Sfl1-mediated repression. These
394 findings not only speak to the critical role of Flo11 in expression of the rough phenotype in the

395 BYx3S cross, but also illustrate how regulation of this gene by multiple pathways determines the
396 phenotypic effects of cryptic variation.
397
398 *GPa and GPb rough segregants utilize different sub-pathways that impact Ras*
399 Excluding the *FLO11* coding region and promoter, all remaining loci act in the presence of only
400 one or two of the GPs. To better understand the highly contextual effects of these cryptic
401 variants, we attempted to resolve detected loci to specific genes. Seven of the GPb-responsive
402 loci detected in this study were previously found to interact with either GPa or other GPs in
403 previous papers, in which they were mapped to individual genes (LEE *et al.* 2016; TAYLOR *et al.*
404 2016) (**Table 1**). These genes encode Ras-regulated transcription factors (*FLO8*, *MSS11*, *MGA1*,
405 *SFL1*), a protein kinase A subunit (*TPK3*), a TOR pathway component (*MDS3)*, and *IRA2*. Two
406 of the loci found in the current study among GPb segregants were located on chromosomes II
407 and XIV, and had never been detected in our past work (**Figure 2, Table 1**). Through genetic
408 engineering experiments, we resolved the chromosome XIV locus to *SRV2* (**Figure S6)**, which
409 encodes a post-translational activator of adenylate cyclase, another component of the Ras
410 pathway. At the chromosome II locus, the most likely candidate is *SRB6*, which encodes an
411 essential subunit of the RNA polymerase II mediator complex. Previously, we showed that
412 mutations disrupting other mediator components, Srb10 and Srb11 (also known as Ssn3 and
413 Ssn8, respectively), can induce the rough phenotype by interacting with a subset of the alleles
414 that have a phenotypic effect in the presence of GPa (TAYLOR *et al.* 2016) (**Note S2**).
415
416 Although loci interacting with GPa and GPb primarily act through the Ras pathway and many
417 loci interact with both GPs, some are specific to one or the other. Among these are *TRR1* and
418 *END3*, which were detected in GPa segregants but not GPb segregants, as well as *MDS3*, *SRV2*,
419 and *TPK3*, which were identified in GPb segregants but not GPa segregants. Notably, these
420 genes play a role in activating the Ras pathway through oxidative stress and actin organization
421 (**Figure 6**). Actin cytoskeleton stability is required for cell polarity and yeast-adhesion traits, and
422 is regulated in part by the effects of End3 and Srv2 on Ras signaling (DU AND AYSCOUGH 2009).
423 This process results in increased production of reactive oxygen species (ROS) through the
424 activity of Tpk3 (GOURLAY AND AYSCOUGH 2006). ROS accumulation is then influenced by
425 Trr1 (CHARIZANIS *et al.* 1999) and the TOR pathway, of which Mds3 is a component. Together,

426 Mds3, Srv2, and Tpk3 form a well-described sub-pathway that influences Ras activity
427 (GOURLAY AND AYSCOUGH 2006; DU AND AYSCOUGH 2009). Thus, although different loci cause
428 the trait in the presence of GPa and GPb, they appear to reflect different sub-pathways affecting
429 the same cellular processes, which ultimately impact how Ras-regulated transcription factors
430 influence *FLO11* activity.
431
432 *Cryptic variation in several pathways underlies the phenotype in GPc segregants*
433 Lastly, we sought to determine the genes harboring cryptic variants that interact with GPc.
434 Regarding GPc, 3 of the 10 identified loci also interact with either GPa or b. These loci
435 correspond to *END3*, *IRA2*, and a locus on chromosome XII. For chromosome XII and the
436 remaining seven loci, we identified likely candidate genes based on our highly resolved genetic
437 mapping data and publicly available research on these genes' functions and phenotypic effects.
438 One of these loci corresponds to *IRA1*, a Ras negative regulator and paralog of *IRA2* (LEE *et al.*
439 2016) (**Table 1**). To obtain additional support for the remaining loci, we performed gene
440 deletions and allele replacements on the identified candidate genes in a GPc rough segregant, and
441 determined the resulting effects on the colony trait (**Methods**). Knock out and replacement of
442 *PMD1, GPA1, SDC25, GPI8, GCN4,* and *YPT6* resulted in loss of the phenotype (**Figure S7a**
443 and **b**), implying that these six genes play positive roles in the phenotype's regulation. In
444 contrast, deletion of *DIG1*, a protein that directly inhibits the transcriptional activator Ste12,
445 enhanced the trait's expression (**Figure S7a** and **c**). These results suggest that genetic variation
446 in these genes plays a causal role in the rough phenotype.
447
448 Detection of *END3*, *IRA1*, and *IRA2* in the absence of Flo8 and Sfl1 indicates that Ras still
449 contributes to *FLO11* regulation in the absence of these transcription factors, possibly by
450 impacting the activities of other transcription factors (ESTRUCH 2000). The remaining loci
451 implicate alternative signaling pathways as being responsible for trait expression in GPc
452 segregants. The genes we identified in the other intervals as having a phenotypic effect when
453 knocked out were a TOR pathway component and *MDS3* paralog (*PMD1*), members of the MAP
454 kinase (MAPK) signaling cascade (*DIG1*, *GPA1*) (METODIEV *et al.* 2002), an environmentally-
455 responsive transcriptional activator (*GCN4),* a Rab GTPase that influences Ras (COSTANZO *et al.*
456 2010), MAPK (COSTANZO *et al.* 2016), and Rim101 (ZHENG *et al.* 2010) signaling *(YPT6),* a

457 stress responsive guanine exchange factor (*SDC25*), and an enzyme that post-translationally
458 modifies proteins to help them anchor into the cell wall (*GPI8*). Of particular note for *GPA1*, the
459 BY strain is known to carry a lab-derived allele that is an expression QTL hotspot (YVERT *et al.*
460 2003), supporting the possibility that *GPA1* allele state might also impact expression of the
461 rough phenotype in the presence of GPc. Additionally, $GPA1^{BY}$ was previously shown to
462 influence other *FLO11*-dependent traits through its downstream transcriptional activator Ste12
463 (MATSUI *et al.* 2015). Nearly all of the genes implicated in allowing GPc segregants to express
464 the rough phenotype have the potential to influence, either directly or indirectly, *FLO11*
465 transcription. The lone exception is *GPI8*, which could still influence Flo11 at the protein level,
466 as it is responsible for adding glycosylphosphatidylinositol (GPI) anchors to new proteins and
467 could affect Flo11's binding to the cell surface (BENGHEZAL *et al.* 1996). These results show the
468 abundant cryptic variation that impacts regulation of *FLO11* and the rough phenotype.
469
470 **Discussion**
471 To better understand the extent of cryptic variation within a population, as well as the
472 mechanisms regulating this cryptic variation, we comprehensively determined the genetic basis
473 of a model phenotype that is only expressed in the presence of particular GPs. By doing this, we
474 identified 21 loci harboring cryptic variants that can contribute to the rough phenotype. Notably,
475 all but one of these loci show phenotypic effects that depend on the GP that is present. In
476 addition, because the cryptic variants that are uncovered by each GP vary in their degree of
477 epistasis with each other and interaction with the environment, we find that the trait's genetic
478 architecture significantly differs across the GPs. This results in a multitude of genotype,
479 environment, and phenotype relationships that produce the same trait.
480
481 Given the detailed understanding that we have obtained for the colony morphology system, a
482 major question is, how might our findings relate to other traits and species? One major insight
483 from our study that may apply to other systems is that most, if not nearly all, of our findings
484 connect to the transcriptional regulation of a single key gene, *FLO11*. Supporting this point, the
485 single locus common to all three GPs is the coding region of *FLO11*, suggesting that regulation
486 of Flo11 levels and stability is the central determinant of the rough phenotype's expression
487 across GPs, combinations of segregating loci, and temperature. Bolstering the importance of

488 variability in *FLO11* regulation to our findings, most of the loci that exhibit phenotypic effects
489 influence *FLO11* in *trans.* Moreover, many of these loci are only visible in the absence of a
490 repressive Rim101 binding site in the *FLO11* promoter. These results highlight the potentially
491 important role that *cis*-regulatory polymorphisms can play in enabling *trans*-regulatory
492 polymorphisms to exert phenotypic effects. Indeed, *cis*-regulatory polymorphisms have been
493 shown to modify the effects of *trans* variants in other systems (REDDY *et al.* 2012; WONG *et al.*
494 2017), thereby altering how genetic differences impact traits (PAYNE AND WAGNER 2014).
495
496 In addition to showing that different GPs enable distinct cryptic variants to have phenotypic
497 effects, we also demonstrated that the GPs modify how genotype-environment interactions
498 impact the trait. For both GPa and GPb, we find that certain combinations of epistatic alleles
499 both enable expression of the trait at a permissive temperature of 21°C, as well as provide the
500 genetic potential for the phenotype at higher temperatures. However, different loci and multi-
501 locus genotypes enable GPa and GPb segregants to express the trait at higher temperatures. In
502 contrast, GPc individuals are unable to express the phenotype at temperatures above 21°C,
503 implying their potential to express the trait across environments is constrained. These findings
504 support the concept that not all genotypes specifying the same trait possess comparable
505 phenotypic robustness (WAGNER 2012; PAYNE *et al.* 2014; PFENNIG AND EHRENREICH 2014;
506 SIEGAL AND LEU 2014; EHRENREICH AND PFENNIG 2016). In the case of the rough phenotype, our
507 genetic mapping and genetic engineering results imply that differences in phenotypic robustness
508 relate to changes in signaling and transcription factor activity across multiple pathways and sub-
509 pathways influencing *FLO11*.
510
511 Furthermore, whereas the rough phenotype in GPa and GPb segregants involves complex
512 epistatic effects, we find no evidence for epistasis in GPc rough segregants. This implies that by
513 eliminating Flo8 and Sfl1, the main transcriptional regulators of *FLO11*, we not only uncovered
514 a different set of cryptic variants but also converted the trait's genetic architecture from mainly
515 epistatic to additive. Perhaps this additivity at the phenotypic level reflects the cumulative effect
516 of multiple pathways influencing *FLO11* expression at the molecular level. While the majority of
517 the loci that interact with GPa and GPb correspond to components of the Ras pathway, many of
518 the loci found in the presence of GPc are involved in other signaling pathways that may have

519 compensatory functions (**Figure 6**). This is consistent with the idea that eliminating Flo8 and
520 Sfl1 might enable other pathways to play a stronger role in expression of *FLO11* and the rough
521 colony phenotype.
522
523 Although not an explicit focus of the current manuscript, our results also provide valuable
524 insights into genetic background effects, the phenomenon in which GPs show different
525 phenotypic effects in distinct individuals (NADEAU 2001; CHANDLER *et al.* 2013; EHRENREICH
526 2017). A number of recent studies have shown that these background effects often result from
527 higher-order epistasis between a GP and multiple segregating loci (CHANDLER *et al.* 2014;
528 TAYLOR AND EHRENREICH 2014; MIOTTO *et al.* 2015; TAYLOR AND EHRENREICH 2015b; TAYLOR
529 AND EHRENREICH 2015a; LEE *et al.* 2016; KUZMIN *et al.* 2018; MULLIS *et al.* 2018). Despite
530 supporting an important role for higher-order epistasis in background effects, our current work,
531 in particular on GPc, also shows that background effects can have much simpler underpinnings.
532 In the context of GPc, identified loci each show pairwise epistasis with the GP, but exhibit no
533 epistasis with each other and thus appear to act additively. These differences in quantitative
534 genetic architecture again appear to tie back to *FLO11* regulation, consistent with theoretical
535 work suggesting that how GPs perturb gene regulatory networks can impact whether loci show
536 additive or epistatic effects (GJUVSLAND *et al.* 2007).
537
538 In summary, our study demonstrates the large amount of cryptic variation that can underlie a
539 single phenotype and strongly implicates complex changes in transcription as a mechanism
540 regulating this cryptic variation. Further, these results suggest a broader insight into the genetic
541 architecture of complex traits. Although it has long been known that traits can vary in genetic
542 architecture depending on the different populations and environments in which they are
543 measured, our findings illustrate that even the same phenotype examined within a single
544 population can exhibit a spectrum of genetic architectures. As shown by the GPs in our study,
545 which of these architectures is visible may depend on only one or two alleles that modify how
546 the rest of the genetic variation in the population behaves. This suggests that characterizing the
547 molecular mechanisms that shape genetic architecture will be an important step in improving our
548 basic understanding of the relationship between genotype and phenotype.
549

Acknowledgments

We thank Joseph Hale, Takeshi Matsui, Martin Mullis, Joann Phan, Rachel Schell, Fabian Seidl, Matthew Taylor, and Gleidson Teixeira for valuable input regarding this project or manuscript. We also thank the anonymous reviewers of this paper for their helpful feedback. The work reported in this paper was supported by grant R01GM110255 from the National Institutes of Health to I.M.E, a fellowship from the Alfred P. Sloan Foundation to I.M.E, and a fellowship from the São Paulo Research Foundation to A.L.V.C.

Data Availability

All sequencing data from this project is available through the National Center for Biotechnology Information (NCBI) Short Read Archive. Data can be accessed under Bioproject ID PRJNA503265, with Biosample accessions SAMN10356503-SAMN10357319.

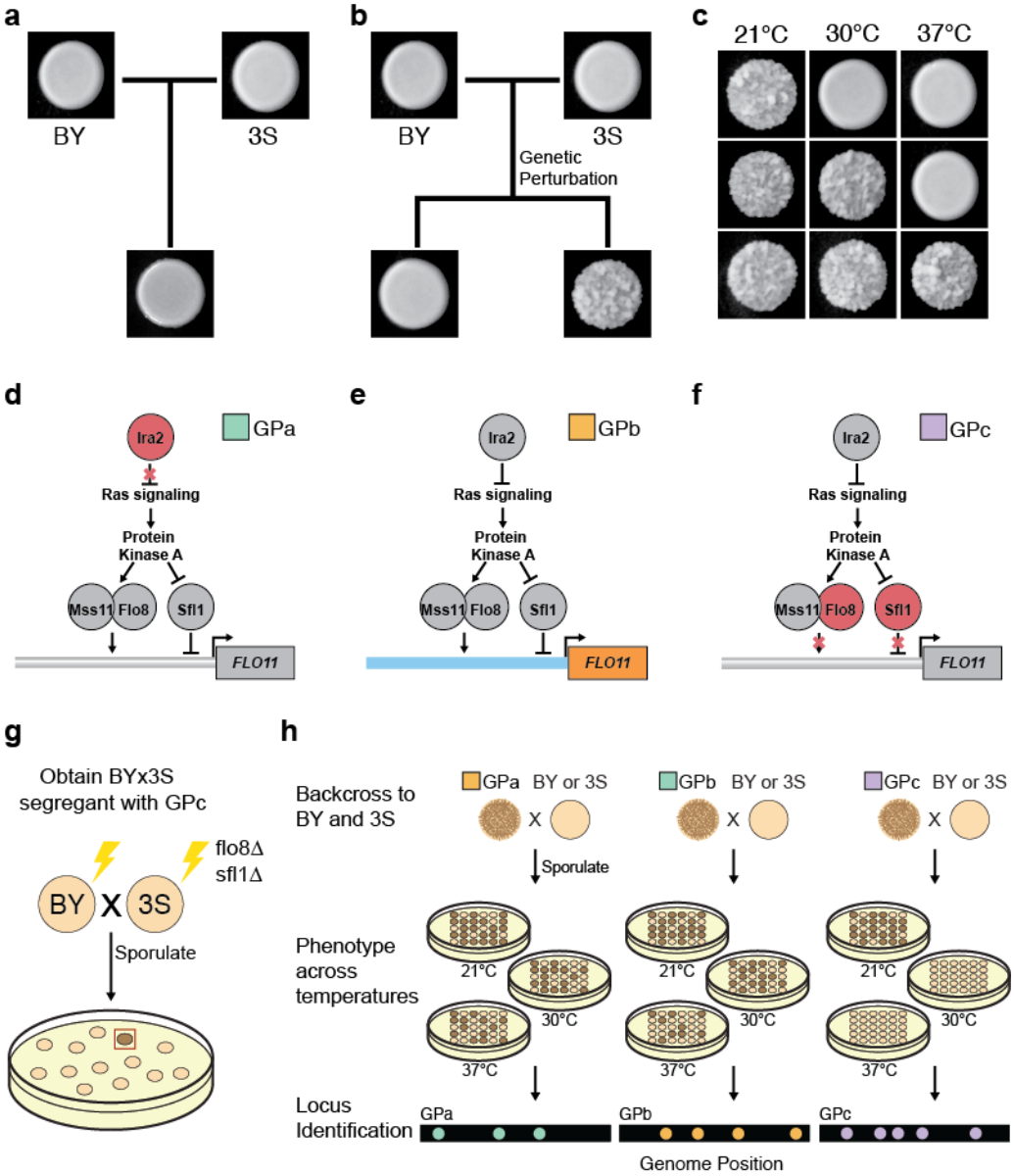


Figure 1: Different genetic perturbations ('GPs') interact with cryptic variation to cause rough morphology. (a) The BY and 3S strains, as well as their cross progeny, form 'smooth' colonies. **(b)** Certain GPs can cause BYx3S segregants to express an alternative 'rough' colony phenotype. **(c)** Expression of the colony trait can be temperature-sensitive, with segregating loci causing individuals form rough colonies at 21°C only, 21 and 30°C, or 21, 30, and 37°C. **(d)** GPa is a loss-of-function mutation in *IRA2*, which encodes a Ras negative regulator. **(e)** GPb results from recombination between the BY (blue) *FLO11* promoter and 3S (orange) *FLO11* coding region. **(f)** GPc is a double knockout of *FLO8* and *SFL1*, which encode the main Ras-dependent transcriptional regulators of *FLO11*. **(g)** GPc was genetically engineered in both BY and 3S, and

572 a rough BYx3S segregant was obtained, as indicated by a brown colony in the illustration. (**h**)
573 Genetic mapping populations were generated by backcrossing a rough segregant with each
574 genetic perturbation to BY and 3S, and screening for the trait at 21, 30, and 37°C. Loci
575 associated with the trait were identified in each backcross for each GP-temperature combination.

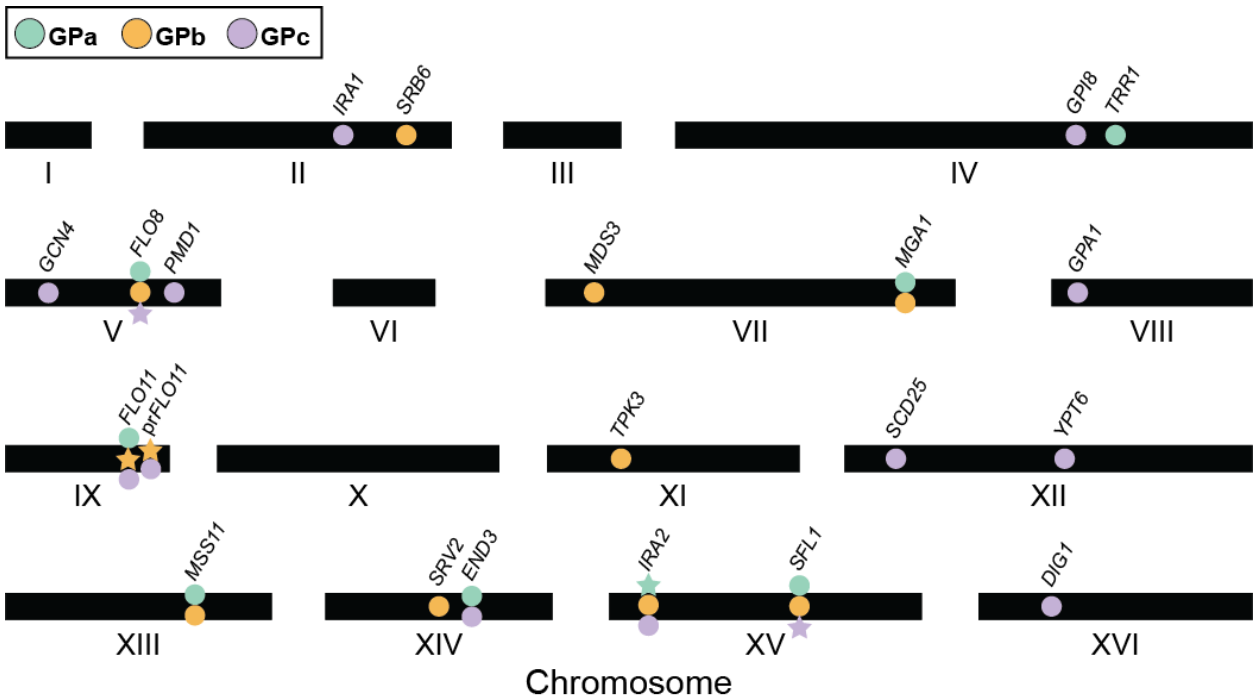
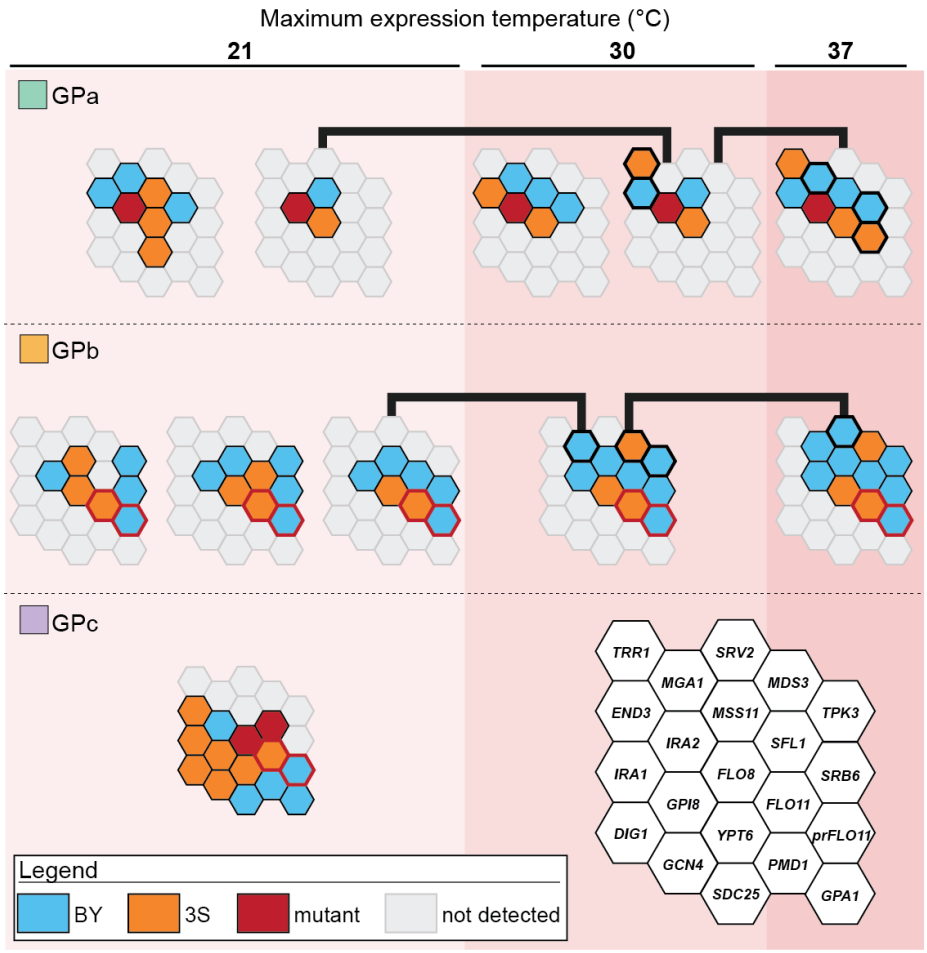


Figure 2: Loci identified in the presence of the GPs. Twenty-one loci were identified in total. Circles indicate cryptic variants, whereas stars indicate GPs. In some instances, the same gene that possesses a GP in one population harbors cryptic variation in another population.



580
581
582 **Figure 3: Loci detected across GPs and temperature sensitivity classes.** (**a**) Loci interacting
583 with GPa and GPb at 21 and 30°C act in different multi-locus genotypes. As shown with black
584 lines, certain temperature-sensitive genotypes provide the potential for additional alleles to
585 enable trait expression at higher temperatures. These additional alleles are designated with bold
586 outlines. In the legend, 'mutant' refers to the *de novo* or induced mutations present involved in
587 each GP. Additionally, the *prFLO11* recombination (GPb) is indicated with a red outline.

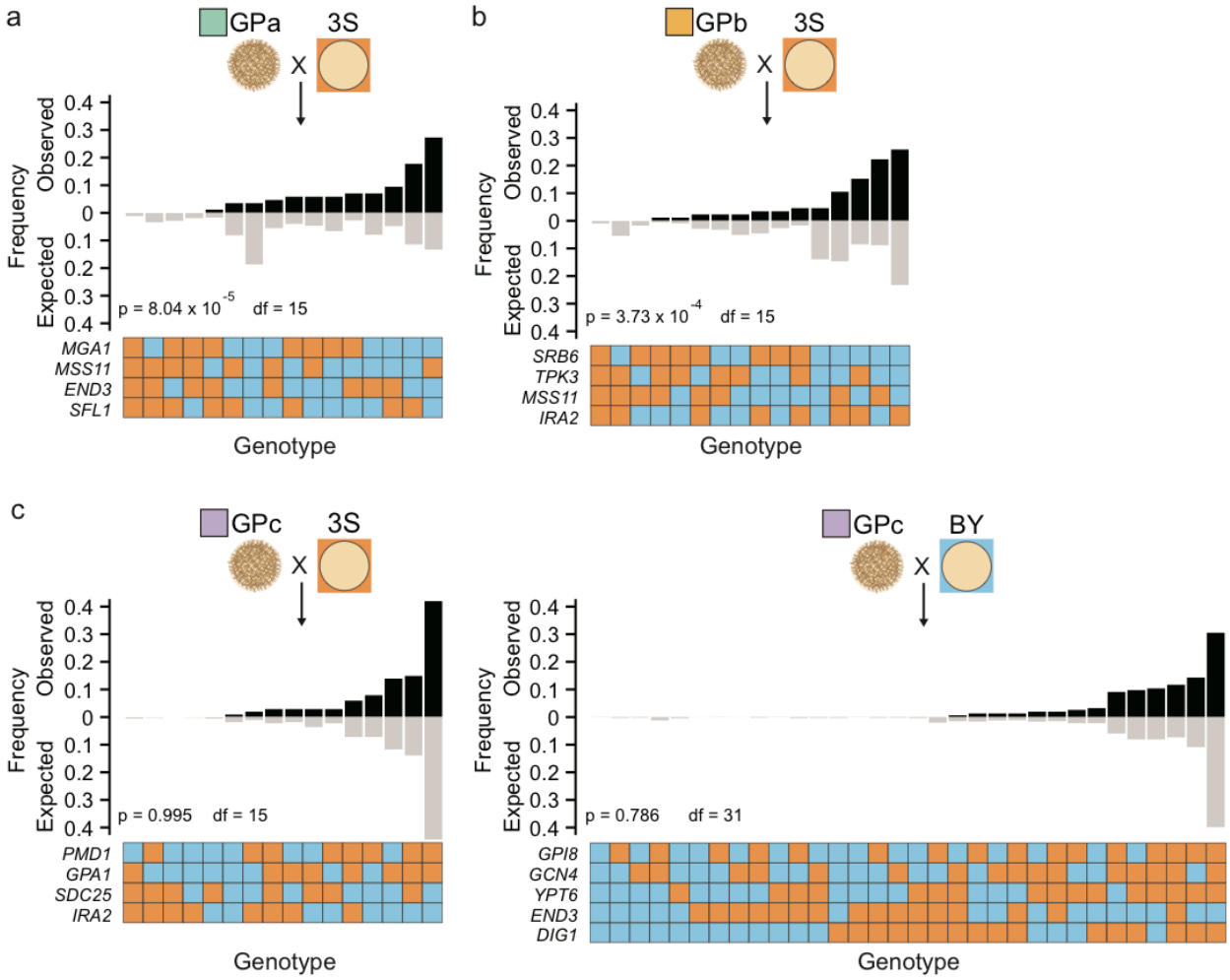


Figure 4: Loci exhibit epistasis or additivity depending on GP. (a) For GPa individuals expressing the trait at only 21°C, observed multi-locus genotype frequencies at the four loci segregating in the 3S backcross do not reflect expected frequencies. This can occur if certain alleles exhibit epistasis and co-segregate in a non-random manner. **(b)** Observed and expected genotype frequencies based on four loci in GPb individuals expressing the trait at only 21°C. **(c)** Observed genotype frequencies in GPc individuals closely match expected frequencies, suggesting involved loci act in an additive manner. In all cases, expected genotype frequencies are computed as the product of the frequency of the involved alleles. P-values correspond to results of a χ^2 test with degrees of freedom indicated in the figure. Loci for which only a single allele is present in their respective population were excluded from analysis, and backcross populations in which only one multi-locus genotype was present are not shown. Only first-generation backcross segregants were used to generate this figure.

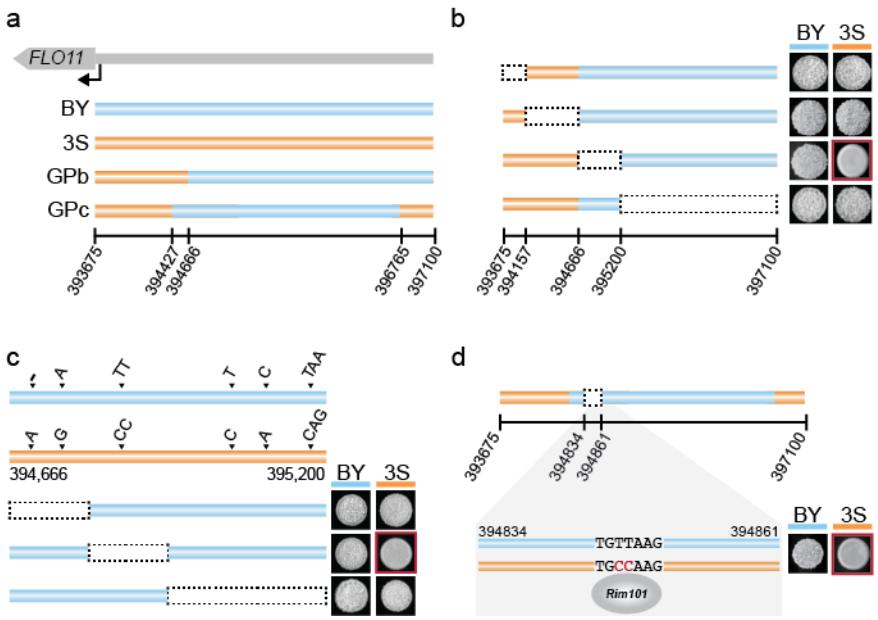
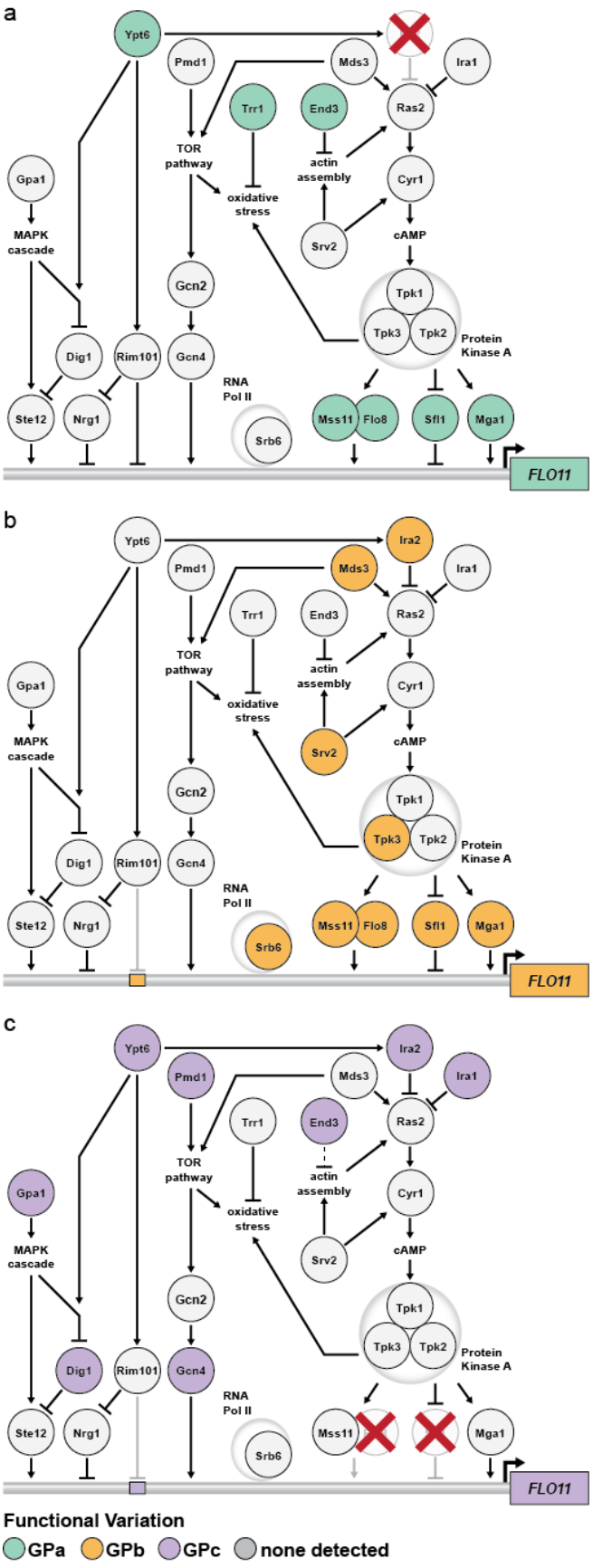


Figure 5: A transcription factor binding site polymorphism is required for GPb and GPc rough segregants to express the trait. (a) The GPb and GPc *FLO11* promoters contain portions of both the BY (blue) and 3S (orange) versions. **(b)** Genetic engineering experiments identified a 534 nucleotide segment of the BY-derived promoter as being responsible for the trait. The causal replacement with the 3S allele leading to loss of the phenotype is outlined in red. **(c)** Engineering of the six polymorphisms within the delimited promoter region revealed that replacement of the TT BY sequence with the 3S CC at position 394,846 to 394,847 allele ablates the phenotype. **(d)** The causal polymorphism results in a Rim101 binding site only present on the 3S promoter. In panels **b** through **d**, red boxes denote allele replacements with phenotypic effects.


a
Ypt6
Pmd1
Mds3
Ira1
Trr1
End3
Ras2
TOR
pathway
actin
assembly
Cyr1
Gpa1
oxidative
stress
MAPK
cascade
Srv2
cAMP
Gcn2
Tpk1
Tpk3
Tpk2
Protein
Kinase A
Dig1
Rim101
Gcn4
RNA
Pol II
Ste12
Nrg1
Srb6
Mss11
Flo8
Sfl1
Mga1
FLO11
b
Ypt6
Ira2
Pmd1
Mds3
Ira1
Trr1
End3
Ras2
TOR
pathway
actin
assembly
Cyr1
Gpa1
oxidative
stress
MAPK
cascade
Srv2
cAMP
Gcn2
Tpk1
Tpk3
Tpk2
Protein
Kinase A
Dig1
Rim101
Gcn4
RNA
Pol II
Ste12
Nrg1
Srb6
Mss11
Flo8
Sfl1
Mga1
FLO11
c
Ypt6
Ira2
Pmd1
Mds3
Ira1
Trr1
End3
Ras2
TOR
pathway
actin
assembly
Cyr1
Gpa1
oxidative
stress
MAPK
cascade
Srv2
cAMP
Gcn2
Tpk1
Tpk3
Tpk2
Protein
Kinase A
Dig1
Rim101
Gcn4
RNA
Pol II
Ste12
Nrg1
Srb6
Mss11
Mga1
FLO11
Functional Variation
GPa
GPb
GPc
none detected

613

614 **Figure 6: GPs unmask cryptic genetic variation in parallel signaling pathways and sub-**
615 **pathways. (a)** Loci that interact with GPa influence *FLO11* regulation by the Ras pathway. **(b)**
616 GPb uncovers loci that primarily act through a different Ras sub-pathway involving Mds3, Srv2,
617 and Tpk3. **(c)** Loci interacting with GPc function in a number of different pathways that are
618 capable of regulating *FLO11* activity. The locations of transcription factor binding sites are not
619 intended to reflect specific positioning along the *FLO11* promoter.

620 **Table 1: Candidate genes and cloned causal genes underlying identified loci**

Genetic Perturbation	Chr.	Gene	Function	Allele	Initial detection
GPc	2	*IRA1*	Ras negative regulator, *IRA2* paralog	3S	Taylor et al., 2016
GPb	2	*SRB6**	RNA Pol II subunit	BY	this paper
GPc	4	*GPI8*	GPI transamidase complex subunit	3S	this paper
GPa	4	*TRR1*	redox state regulator	3S	Taylor and Ehrenreich, 2014
GPc	5	*GCN4*	stress response transcription factor	3S	this paper
GPa,b	5	*FLO8*	Ras activated transcription factor	3S	Taylor and Ehrenreich, 2014
GPc	5	*PMD1*	growth regulator; *MDS3* paralog	BY	this paper
GPb	7	*MDS3*	TOR pathway growth regulator	3S	Taylor et al., 2016
GPa,b	7	*MGA1*	transcriptional activator	BY	Taylor and Ehrenreich, 2015
GPc	8	*GPA1*	G-protein	BY	this paper
GPa,b,c	9	*FLO11*	cell surfaace glycoprotein	3S	Lee et al., 2016
GPb,c	9	*prFLO11*	*FLO11* promoter	BY	Taylor et al., 2016
GPb	11	*TPK3*	Protein Kinase A subunit	BY	Taylor et al., 2016
GPc	12	*SDC25*	Ras GEF	BY	this paper
GPa,c	12	*YPT6*	Rab GTPase	3S	Lee et al., 2016
GPa,b	13	*MSS11*	Ras activated transcription factor	BY,3S	Taylor and Ehrenreich, 2014
GPb	14	*SRV2*	activator of adenylate cyclase	BY	this paper
GPa,c	14	*END3*	cell wall morphogenesis	BY, 3S	Taylor and Ehrenreich, 2014
GPb,c	15	*IRA2*	Ras negative regulator	BY	Taylor et al., 2016
GPa,b	15	*SFL1*	Ras inactivated transcription factor	BY	Taylor and Ehrenreich, 2015
GPc	16	*DIG1*	MAPK-regulated transcription factor	3S	this paper

621 * candidate gene

622 **References**

623 Atwell, S., Y. S. Huang, B. J. Vilhjalmsson, G. Willems, M. Horton *et al.*, 2010 Genome-wide
624 association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465**:**
625 627-631.

626 Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic
627 analysis of complex traits in the emerging Collaborative Cross. Genome Res 21**:** 1213-
628 1222.

629 Barrales, R. R., J. Jimenez and J. I. Ibeas, 2008 Identification of novel activation mechanisms for
630 FLO11 regulation in Saccharomyces cerevisiae. Genetics 178**:** 145-156.

631 Benghezal, M., A. Benachour, S. Rusconi, M. Aebi and A. Conzelmann, 1996 Yeast Gpi8p is
632 essential for GPI anchor attachment onto proteins. EMBO J 15**:** 6575-6583.

633 Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and
634 powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B
635 57**:** 289-300.

636 Bergman, A., and M. L. Siegal, 2003 Evolutionary capacitance as a general feature of complex
637 gene networks. Nature 424**:** 549-552.

638 Bloom, J. S., I. M. Ehrenreich, W. T. Loo, T. L. Lite and L. Kruglyak, 2013 Finding the sources
639 of missing heritability in a yeast cross. Nature 494**:** 234-237.

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

642 Chandler, C. H., S. Chari and I. Dworkin, 2013 Does your gene need a background check? How
643 genetic background impacts the analysis of mutations, genes, and evolution. Trends
644 Genet 29**:** 358-366.

645 Chandler, C. H., S. Chari, D. Tack and I. Dworkin, 2014 Causes and consequences of genetic
646 background effects illuminated by integrative genomic analysis. Genetics 196**:** 1321-
647 1336.

648 Charizanis, C., H. Juhnke, B. Krems and K. D. Entian, 1999 The oxidative stress response
649 mediated via Pos9/Skn7 is negatively regulated by the Ras/PKA pathway in
650 Saccharomyces cerevisiae. Mol Gen Genet 261**:** 740-752.

651 Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear *et al.*, 2010 The genetic
652 landscape of a cell. Science 327**:** 425-431.

653 Costanzo, M., B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons *et al.*, 2016 A global
654 genetic interaction network maps a wiring diagram of cellular function. Science 353.
655 Du, W., and K. R. Ayscough, 2009 Methyl beta-cyclodextrin reduces accumulation of reactive
656 oxygen species and cell death in yeast. Free Radic Biol Med 46**:** 1478-1487.
657 Dworkin, I., A. Palsson, K. Birdsall and G. Gibson, 2003 Evidence that Egfr contributes to
658 cryptic genetic variation for photoreceptor determination in natural populations of
659 Drosophila melanogaster. Curr Biol 13**:** 1888-1893.
660 Ehrenreich, I. M., 2017 Epistasis: Searching for Interacting Genetic Variants Using Crosses.
661 Genetics 206**:** 531-535.
662 Ehrenreich, I. M., and D. W. Pfennig, 2016 Genetic assimilation: a review of its potential
663 proximate causes and evolutionary consequences. Ann Bot 117**:** 769-779.
664 Estruch, F., 2000 Stress-controlled transcription factors, stress-induced genes and stress tolerance
665 in budding yeast. FEMS Microbiol Rev 24**:** 469-486.
666 Fidalgo, M., R. R. Barrales, J. I. Ibeas and J. Jimenez, 2006 Adaptive evolution by mutations in
667 the FLO11 gene. Proc Natl Acad Sci U S A 103**:** 11228-11233.
668 Fidalgo, M., R. R. Barrales and J. Jimenez, 2008 Coding repeat instability in the FLO11 gene of
669 Saccharomyces yeasts. Yeast 25**:** 879-889.
670 Geiler-Samerotte, K. A., Y. O. Zhu, B. E. Goulet, D. W. Hall and M. L. Siegal, 2016 Selection
671 Transforms the Landscape of Genetic Variation Interacting with Hsp90. PLoS Biol 14**:**
672 e2000465.
673 Gibson, G., and I. Dworkin, 2004 Uncovering cryptic genetic variation. Nat Rev Genet 5**:** 681-
674 690.
675 Gietz, R. D., and R. A. Woods, 2002 Transformation of yeast by lithium acetate/single-stranded
676 carrier DNA/polyethylene glycol method. Methods Enzymol 350**:** 87-96.
677 Gjuvsland, A. B., B. J. Hayes, S. W. Omholt and O. Carlborg, 2007 Statistical epistasis is a
678 generic feature of gene regulatory networks. Genetics 175**:** 411-420.
679 Gourlay, C. W., and K. R. Ayscough, 2006 Actin-induced hyperactivation of the Ras signaling
680 pathway leads to apoptosis in Saccharomyces cerevisiae. Mol Cell Biol 26**:** 6487-6501.
681 Hope, E. A., and M. J. Dunham, 2014 Ploidy-regulated variation in biofilm-related phenotypes in
682 natural isolates of Saccharomyces cerevisiae. G3 (Bethesda) 4**:** 1773-1786.

683 Jarosz, D. F., and S. Lindquist, 2010 Hsp90 and environmental stress transform the adaptive
684 value of natural genetic variation. Science 330: 1820-1824.
685 Kuchin, S., V. K. Vyas and M. Carlson, 2002 Snf1 protein kinase and the repressors Nrg1 and
686 Nrg2 regulate FLO11, haploid invasive growth, and diploid pseudohyphal differentiation.
687 Mol Cell Biol 22: 3994-4000.
688 Kuzmin, E., B. VanderSluis, W. Wang, G. Tan, R. Deshpande *et al.*, 2018 Systematic analysis of
689 complex genetic interactions. Science 360.
690 Lamb, T. M., and A. P. Mitchell, 2003 The transcription factor Rim101p governs ion tolerance
691 and cell differentiation by direct repression of the regulatory genes NRG1 and SMP1 in
692 Saccharomyces cerevisiae. Mol Cell Biol 23: 677-686.
693 Laughery, M. F., T. Hunter, A. Brown, J. Hoopes, T. Ostbye *et al.*, 2015 New vectors for simple
694 and streamlined CRISPR-Cas9 genome editing in Saccharomyces cerevisiae. Yeast 32:
695 711-720.
696 Le Rouzic, A., and O. Carlborg, 2008 Evolutionary potential of hidden genetic variation. Trends
697 Ecol Evol 23: 33-37.
698 Lee, J. T., M. B. Taylor, A. Shen and I. M. Ehrenreich, 2016 Multi-locus Genotypes Underlying
699 Temperature Sensitivity in a Mutationally Induced Trait. PLoS Genet 12: e1005929.
700 Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler
701 transform. Bioinformatics 25: 1754-1760.
702 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence
703 Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.
704 Lo, W. S., and A. M. Dranginis, 1996 FLO11, a yeast gene related to the STA genes, encodes a
705 novel cell surface flocculin. J Bacteriol 178: 7144-7151.
706 Mackay, T. F., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The Drosophila
707 melanogaster Genetic Reference Panel. Nature 482: 173-178.
708 Matsui, T., R. Linder, J. Phan, F. Seidl and I. M. Ehrenreich, 2015 Regulatory Rewiring in a
709 Cross Causes Extensive Genetic Heterogeneity. Genetics 201: 769-777.
710 McGuigan, K., and C. M. Sgro, 2009 Evolutionary consequences of cryptic genetic variation.
711 Trends Ecol Evol 24: 305-311.
712 Metodiev, M. V., D. Matheos, M. D. Rose and D. E. Stone, 2002 Regulation of MAPK function
713 by direct interaction with the mating-specific Galpha in yeast. Science 296: 1483-1486.

714 Miotto, O., R. Amato, E. A. Ashley, B. MacInnis, J. Almagro-Garcia *et al.*, 2015 Genetic
715 architecture of artemisinin-resistant Plasmodium falciparum. Nat Genet 47**:** 226-234.
716 Mullis, M. N., T. Matsui, R. Schell, R. Foree and I. M. Ehrenreich, 2018 The complex
717 underpinnings of genetic background effects. Nat Commun 9**:** 3548.
718 Nadeau, J. H., 2001 Modifier genes in mice and humans. Nat Rev Genet 2**:** 165-174.
719 Paaby, A. B., and M. V. Rockman, 2014 Cryptic genetic variation: evolution's hidden substrate.
720 Nat Rev Genet 15**:** 247-258.
721 Payne, J. L., J. H. Moore and A. Wagner, 2014 Robustness, evolvability, and the logic of genetic
722 regulation. Artif Life 20**:** 111-126.
723 Payne, J. L., and A. Wagner, 2014 The robustness and evolvability of transcription factor
724 binding sites. Science 343**:** 875-877.
725 Pfennig, D. W., and I. M. Ehrenreich, 2014 Towards a gene regulatory network perspective on
726 phenotypic plasticity, genetic accommodation and genetic assimilation. Mol Ecol 23**:**
727 4438-4440.
728 Queitsch, C., T. A. Sangster and S. Lindquist, 2002 Hsp90 as a capacitor of phenotypic variation.
729 Nature 417**:** 618-624.
730 Reddy, T. E., J. Gertz, F. Pauli, K. S. Kucera, K. E. Varley *et al.*, 2012 Effects of sequence
731 variation on differential allelic transcription factor occupancy and gene expression.
732 Genome Res 22**:** 860-869.
733 Rutherford, S. L., and S. Lindquist, 1998 Hsp90 as a capacitor for morphological evolution.
734 Nature 396**:** 336-342.
735 Seabold, S., and J. Perktold, 2010 Statsmodels: econometric and statistical modeling with
736 Python. Proc of the 9th Python in Science Conf**:** 57-61.
737 Siegal, M. L., and J. Y. Leu, 2014 On the Nature and Evolutionary Impact of Phenotypic
738 Robustness Mechanisms. Annu Rev Ecol Evol Syst 45**:** 496-517.
739 Taylor, M. B., and I. M. Ehrenreich, 2014 Genetic interactions involving five or more genes
740 contribute to a complex trait in yeast. PLoS Genet 10**:** e1004324.
741 Taylor, M. B., and I. M. Ehrenreich, 2015a Higher-order genetic interactions and their
742 contribution to complex traits. Trends Genet 31**:** 34-40.
743 Taylor, M. B., and I. M. Ehrenreich, 2015b Transcriptional Derepression Uncovers Cryptic
744 Higher-Order Genetic Interactions. PLoS Genet 11**:** e1005606.

745 Taylor, M. B., J. Phan, J. T. Lee, M. McCadden and I. M. Ehrenreich, 2016 Diverse genetic
746 architectures lead to the same cryptic phenotype in a yeast cross. Nat Commun 7**:** 11669.
747 Tong, A. H., M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader *et al.*, 2001 Systematic genetic
748 analysis with ordered arrays of yeast deletion mutants. Science 294**:** 2364-2368.
749 Verstrepen, K. J., A. Jansen, F. Lewitter and G. R. Fink, 2005 Intragenic tandem repeats generate
750 functional variability. Nat Genet 37**:** 986-990.
751 Wach, A., A. Brachat, R. Pohlmann and P. Philippsen, 1994 New heterologous modules for
752 classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 10**:** 1793-
753 1808.
754 Wagner, A., 2012 The role of robustness in phenotypic adaptation and innovation. Proc Biol Sci
755 279**:** 1249-1258.
756 Wong, E. S., B. M. Schmitt, A. Kazachenka, D. Thybert, A. Redmond *et al.*, 2017 Interplay of
757 cis and trans mechanisms driving transcription factor binding and gene expression
758 evolution. Nat Commun 8**:** 1092.
759 Yvert, G., R. B. Brem, J. Whittle, J. M. Akey, E. Foss *et al.*, 2003 Trans-acting regulatory
760 variation in Saccharomyces cerevisiae and the role of transcription factors. Nat Genet 35**:**
761 57-64.
762 Zara, G., S. Zara, C. Pinna, S. Marceddu and M. Budroni, 2009 FLO11 gene length and
763 transcriptional level affect biofilm-forming ability of wild flor strains of Saccharomyces
764 cerevisiae. Microbiology 155**:** 3838-3846.
765 Zheng, J., J. J. Benschop, M. Shales, P. Kemmeren, J. Greenblatt *et al.*, 2010 Epistatic
766 relationships reveal the functional organization of yeast transcription factors. Mol Syst
767 Biol 6**:** 420.
768

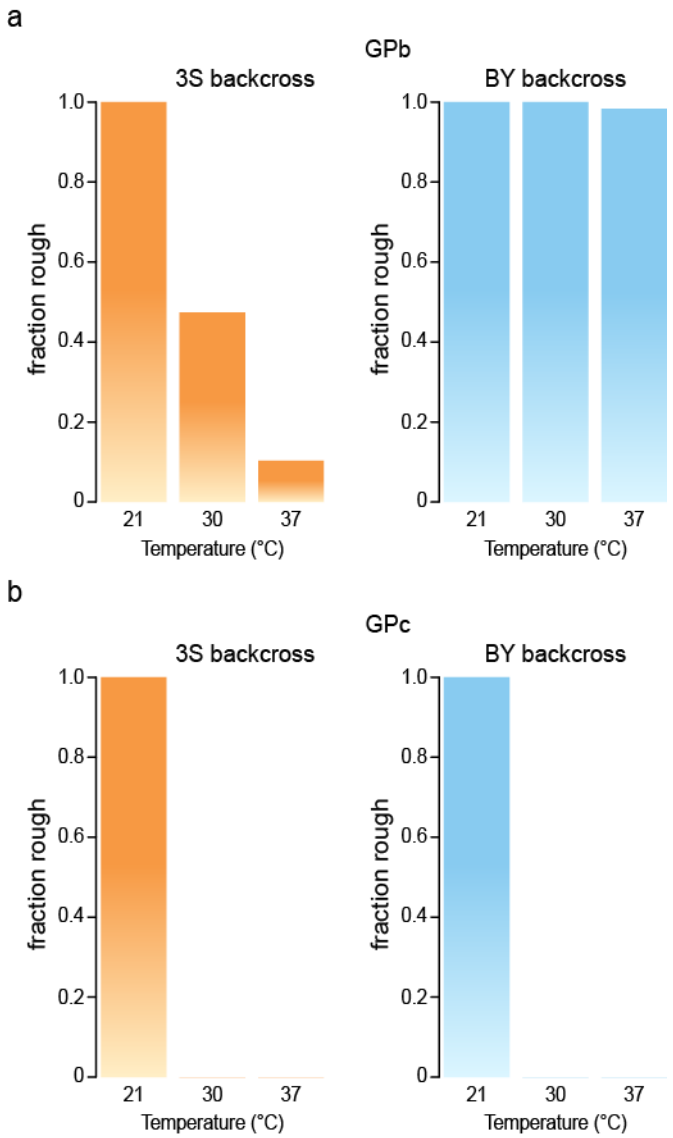


Figure S1: GPb and GPc individuals differ in their temperature sensitivities. (a) A subset of GPb backcross segregants express rough morphology at 30 and 37°C. **(b)** GPc backcross segregants are incapable of expressing the trait at 30 and 37°C. Ninety-six backcross segregants were generated and phenotyped for each GP-backcross parent combination.

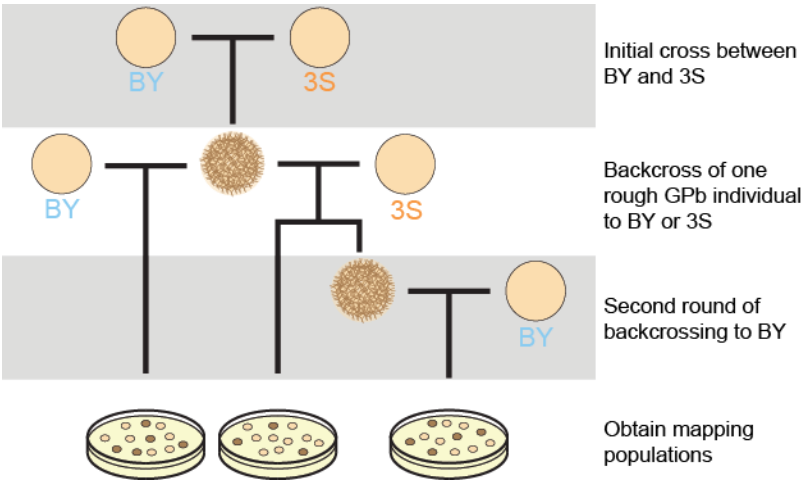


Figure S2: Scheme to generate GPb backcross mapping populations. A rough BYx3S F_2 segregant carrying GPb was backcrossed to BY and 3S. In the case of the cross to 3S, a second-generation backcross to BY was then performed, which was necessary to enrich for segregants that express the rough phenotype at 37°C.

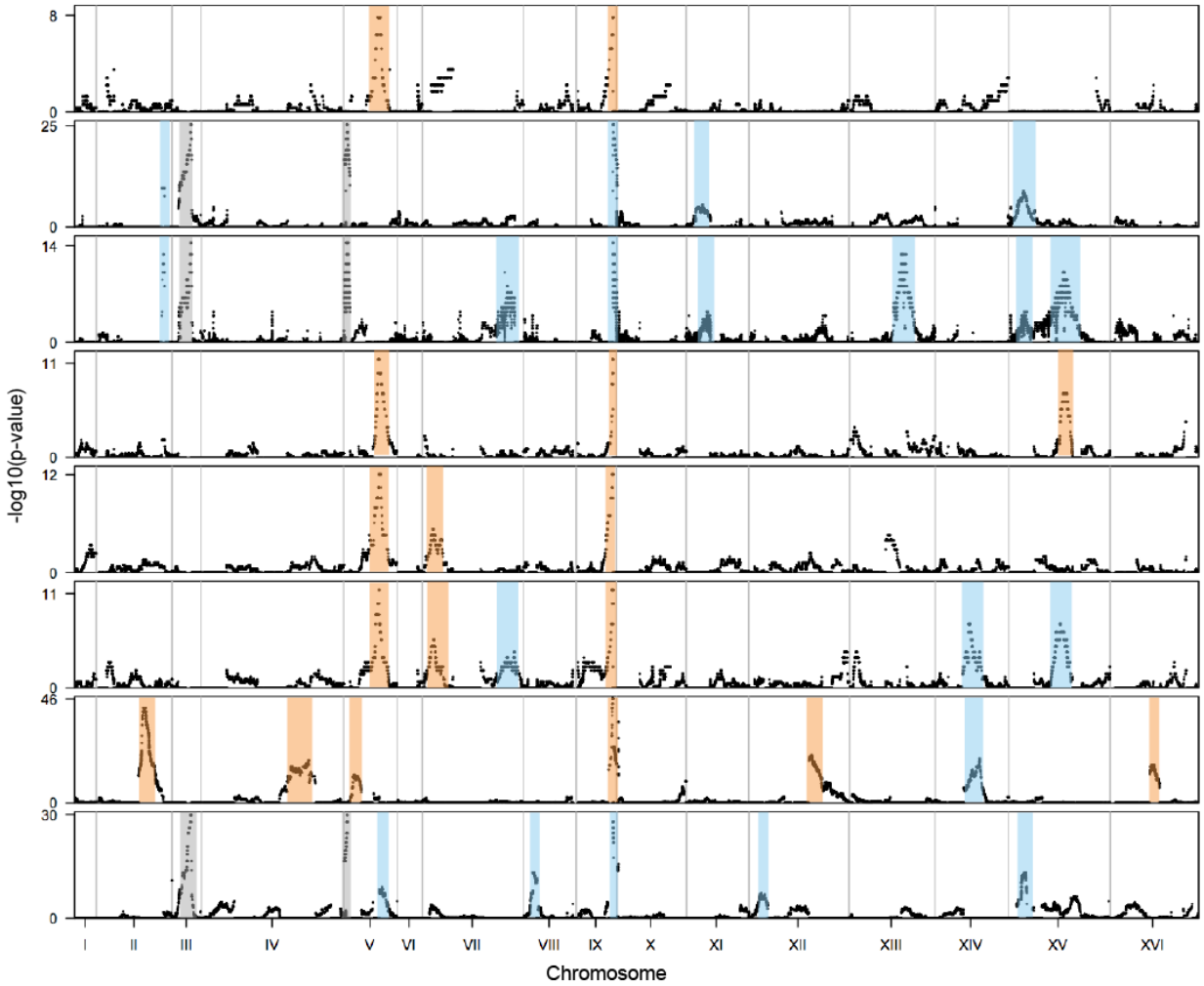


Figure S3: Genetic mapping for loci interacting with GPb and GPc. For each mapping population based on GP, backcross, and temperature sensitivity, a binomial test was performed for each SNP to identify regions of the genome associated with the rough phenotype. Results in are plotted in -$\log_{10}$(p)-values in the following order: GPb (BY backcross, 21, 30, 37°C), GPb (3S backcross, 21°C), GPb (3S backcross, 21, 30°C), GPb (2nd generation backcross, 21°C), GPb (2nd generation backcross, 21, 30°C), GPb (2nd generation backcross, 21, 30, 37°C), GPc (BY backcross, 21°C), GPc (3S backcross, 21°C). Blue bars indicate loci originating from BY, and orange bars indicate loci from 3S. Gray bars denote selectable markers used in generating mapping populations. Details regarding each locus are found in Table S1.

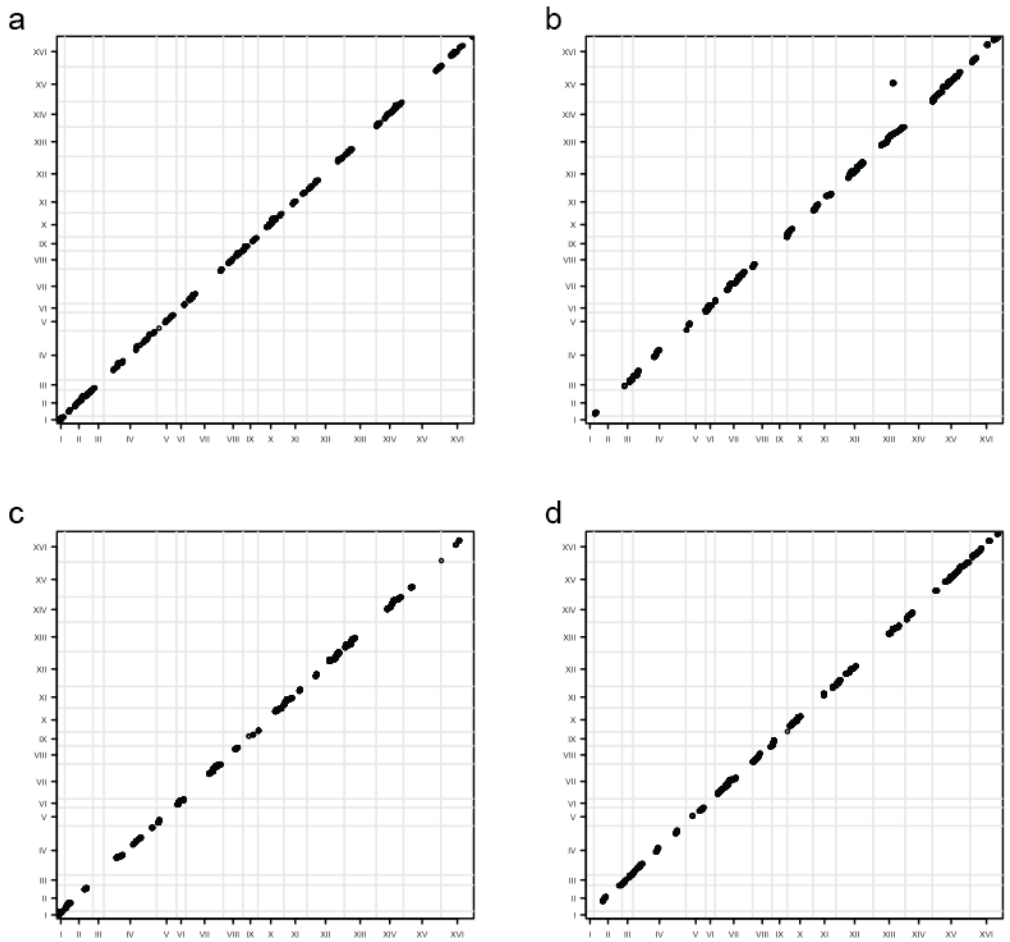


Figure S4: Scan for loci with correlated allele states. Significant p-values for χ^2 tests between each pair of uniquely segregating genomic regions: GPb – **(a)** backcross to BY, **(b)** backcross to 3S, and GPc – **(c)** backcross to BY, **(d)** backcross to 3S. Results are only plotted for the upper triangle. Significant results were only found in the GPb backcross to 3S, identifying an interaction between the genes *MSS11* and *SFL1*.

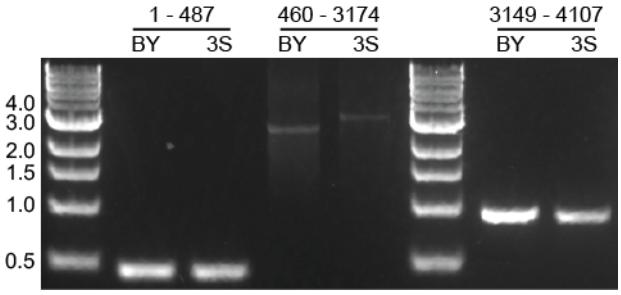


Figure S5: Variation in the *FLO11* coding region. The 3S strain carries a longer *FLO11* allele relative to BY. This is due to a ~500-600 nucleotide length polymorphism in the middle of the gene. Nucleotide positions are provided based on the BY gene sequence.

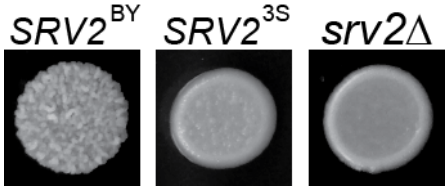


Figure S6: The *SRV2* gene underlies a locus on chromosome XIV that interacts with GPb. Replacement of the BY *SRV2* allele with the 3S version results in loss of the rough phenotype. A similar loss of phenotype is also observed when *SRV2* is deleted.

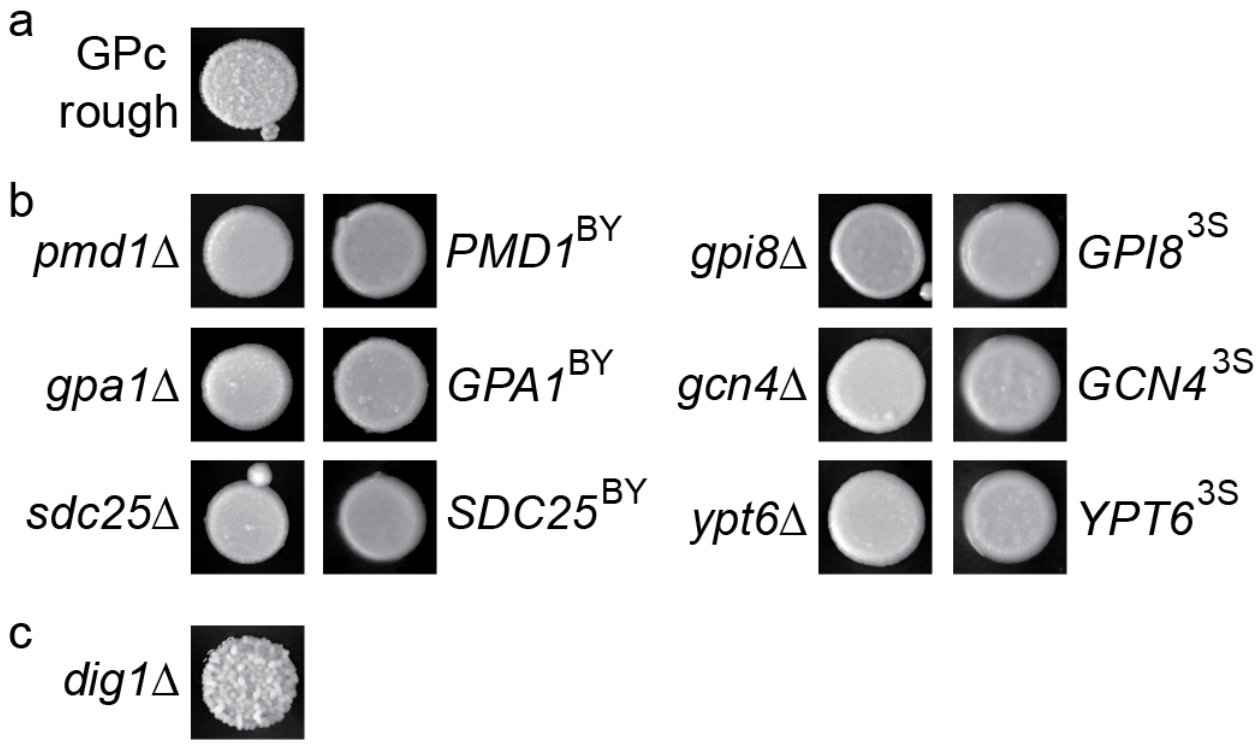


Figure S7: Genetic engineering supports interactions between GPc and identified candidate genes. Complete deletion of the candidate genes in a GPc rough segregant was performed using tailed *kanMX* cassettes. The phenotype of the GPc rough segregant is shown in **(a)**. In **(b)**, we show that deletion of six candidate genes causes loss of the phenotype. We also show that allele replacement with the non-causal allele at these genes results in no phenotypic recovery, implying these genes harbor functional genetic variation. Note, these replacements involved the entire coding region, as well as 100 to 200 bases of promoter sequence and 100 to 200 bases of downstream sequence. Control experiments were performed in parallel and resulted in restoration of the phenotype. In **(c)**, for the seventh candidate gene, deletion enhances the phenotype. Notably, *DIG1* encodes a repressor of the *FLO11* activator Ste12. Thus, the phenotypic enhancement seen upon its deletion is consistent with cryptic variation promoting transcription of *FLO11*.

817 **Note S1:** For GPb individuals expressing the rough phenotype at 21°C, we detected two multi-
818 locus genotypes among segregants from the backcross to 3S. We also identified a third genotype
819 in the second-generation backcross population, involving a different combination of alleles than
820 those from the initial 3S backcross.
821
822 **Note S2:** We deleted each non-essential gene at the GPb chromosome II locus – *RRT2*, *HIS7*,
823 *ARO4*, *MRPS5*, *MTC4*, *SHG1*, and *YBR259W* – and found that none of these deletions caused
824 loss of the rough phenotype. This implies that loss-of-function in a non-essential gene is not
825 responsible for the trait at this locus, and the causal genetic variant is either a gain-of-function
826 polymorphism or is an essential gene. Among the essential genes at this locus, *SRB6* was
827 determined to be the strongest candidate.