Trait specificity mediated by alternative DNA-binding preferences of a single transcription factor in yeast.

Authors: Michael W. Dorrity[1,2], Josh T. Cuperus[1,3], Jolie A. Carlisle[1], Stanley Fields[1,3], Christine Queitsch[1*]

Affiliations:
[1] Department of Genome Sciences, University of Washington, Seattle, WA 98195
[2] Department of Biology, University of Washington, Seattle, WA 98195
[3] Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195
[4] Department of Medicine, University of Washington, Seattle, WA 98195

[*] Correspondence to: queitsch@u.washington.edu

Abstract: In the yeast *Saccharomyces cerevisiae* and other fungal species, the decision to mate or invade relies on a shared transcription factor, Ste12. Specificity toward invasion is attributed to Ste12 binding cooperatively with the co-factor Tec1. Here, we show that single amino acid mutations in the DNA-binding domain of Ste12 suffice to shift the preference of *S. cerevisiae* cells toward either mating or invasion, independent of Tec1. Mutations that result in a shift toward one or the other trait define two distinct regions of this domain, implying an alternative DNA-binding mode for each trait. Using binding site library selections, we show that Ste12 variants that shift trait preference indeed favor alternative DNA-binding sites. We identify temperature-responsive Ste12 variants that promote increased invasion of *S. cerevisiae* only at high temperature, behavior more typical of a fungal pathogen. Lastly, we show that rare Ste12 mutations confer Hsp90-dependent mating, even though wild-type Ste12 is not an Hsp90 client. In summary, we show that flexibility at the level of Ste12 DNA-binding mediates environmentally-sensitive control of two complex traits.

In the yeast *S. cerevisiae,* mating is initiated by pheromone recognition, which activates an evolutionarily conserved G-protein-coupled MAPK pathway (*1*) (Fig. 1A). Invasion is initiated in response to changes in temperature and nutrient availability. The shared MAPKKK Ste11, a known client of Hsp90 (*2*), likely contributes to the environmental sensitivity of both traits. Specificity between mating and invasion pathways is generated at multiple levels. For example, a mating-specific scaffold protein, Ste5, guides kinase signal transduction to activate mating (*3*). Two MAP kinases, Fus3 and Kss1, have overlapping functions in mating but opposing functions in invasion (*4*). At the level of transcriptional control, the two pathways converge on Ste12, which interacts with cofactors to activate either mating or invasion. For mating, Ste12 can bind at pheromone-responsive genes as a homodimer or with the cofactors Mcm1 and Matα1 (*5–7*). For invasion, Ste12 has been proposed to bind cooperatively with Tec1 at invasion genes (*8–10*); an alternative model, however, posits that a complex of Ste12 and Tec1 act solely through Tec1 binding sites (*10*). However, Tec1 is either is not present or not required for invasion in many fungal species, though nearly all species contain Ste12 (Fig. 1- Suppl. 1). We speculated that the highly conserved Ste12 DNA binding domain contributes to the choice between mating and invasion, possibly through alternative DNA binding modes.

Results

To examine the interaction between mating and invasion, we used deep mutational scanning of a segment of Ste12's DNA-binding domain. We generated ~20,000 protein variants within 33 amino acids of this domain, including single, double, and higher order amino acid mutants, and subjected yeast cells carrying this variant library to selection for either mating or invasion (Fig. 1B, Fig. 1 – Suppl. 2, Fig. 1 - Suppl. 3). As a control, we employed selection for a third trait, response to osmotic stress, which shares upstream pathway components but does not involve Ste12. We expected that most *STE12* mutations would affect mating and invasion similarly, because the Ste12 DNA-binding domain is highly conserved among fungi and Ste12 transcriptional activity is required for both traits. Indeed, we found positions in which almost any amino acid substitution was highly deleterious to both mating and invasion (Fig. 1C, 1D, Fig. 1 – Suppl. 3). Conservation is only partially predictive for these results as some of these positions are invariant among fungi (W156, C144), whereas others are not (K149, Q151, K152). Unexpectedly, our mutational analysis also revealed 'separation-of-function' mutations, which primarily reduced either mating or invasion. Substitutions with mostly deleterious effects on mating clustered in residues N-terminal to the conserved W156 (designated region I), while substitutions with deleterious effects on invasion appeared more frequently C-terminal to W156 (designated region II). Substitutions that increased mating were rare. By contrast, substitutions at some positions, almost all within region I, increased invasion, and these substitutions also decreased mating. Thus, mutations in the Ste12 DNA-binding domain can impose preference for mating or invasion rather than similarly affecting both traits.

As these results suggested that the Ste12 DNA-binding domain itself can modulate the mating versus invasion preference, we further explored the apparent structure of mutational effects through double mutant analysis (Fig. 2A). For each pair of positions, we obtained the functional score for mating and for invasion due to every double mutant combination that included this pair. We used these scores to calculate a mean effect on the two traits for each set of double mutant combinations (Fig. 2A, example pair indicated). This analysis confirmed the bipartite arrangement of the mutagenized segment, split by the central W156 (Fig. 2A). Double mutants involving two positions in region I were mostly defective for mating, while those involving two positions in region II were mostly defective for invasion. Double mutants between regions I and II showed variability in mutational effects, with some pairwise combinations increasing mating and others decreasing mating; the effects on invasion, however, for these same combinations did not follow the same pattern. For example, for all variants with a mutation at K150 in region I, the deleterious effect on mating could be rescued slightly, depending on the position mutated in region II (examples highlighted in Fig. 2A, Fig. 2 – Suppl. 2). This pattern was not maintained for invasion, suggesting that K150 affects invasion-specific function differently.

We calculated a "specificity score" for single mutations as the ratio of their effects on mating and invasion (Fig. 2B). The distribution had a strong negative skew, indicating that most single mutations shifted trait preference toward invasion. This distribution, along with the rarity of variants that increased mating above the wild-type level, leads us to speculate that wild-type *S. cerevisiae* Ste12 promotes mating more readily than invasion.

To explore the apparent tradeoff between mating and invasion in more detail, we used the Pareto front concept (*11*). This concept, rooted in engineering and economics, defines all feasible

solutions for optimizing performance in two essential tasks; here, we consider all feasible genotypes that affect mating or invasion. We plotted the single mutation mean positional scores for both traits, identifying positions close to the Pareto front that show increased invasion and deleterious effects on mating. (Fig. 2C, see also Fig. 2 – Suppl. 3). Positions well below the Pareto front, including most prominently W156, had deleterious effects on both traits. These positions are significantly more conserved among fungi, consistent with variation at these sites being disfavored given the constraint on Ste12 to maintain both mating and invasion function in the fungal lineage (Fig. 2D).

Region I or II mutations did not shift the preference for mating or invasion uniformly; rather, effects were position-specific. For example, most mutations at K152 shifted trait preference strongly towards invasion, whereas those at F157 caused more gradual effects (Fig. 2 – Suppl. 4). Positions with extreme specificity scores clustered according to trait preference (Fig. 2E, blue and grey shading). Eliminating positive charges in region I, which may reduce the affinity of Ste12 for DNA, shifted trait preference towards invasion. We found little natural variation affecting charge in region I and II, but detected such variation outside of the mutagenized region (Fig. 2 – Suppl. 1). For example, the Ste12 DNA-binding domain of the invasive pathogen *Cryptococcus gattii* (*12*) contains two additional positive residues immediately C-terminal to region II (Fig. 2E). Introducing these *C. gattii*-specific residues (S177K, Q180R; denoted SKQR) into the *S. cerevisiae* Ste12 DNA-binding domain yielded a dominant invasion phenotype (Fig. 2E), and decreased mating (Fig. 2 – Suppl. 5).

Tec1 is thought to play an essential role in invasion by activating invasion genes, with Ste12 either binding directly to DNA cooperatively with Tec1 or indirectly as part of a complex with Tec1 (*9*, *10*). However, in contrast to both models, SKQR's invasive phenotype was independent of Tec1, as was the even stronger invasion phenotype due to the K150A variant (Fig. 2E). Thus, a single mutation in the Ste12 DNA-binding domain is sufficient to promote Tec1-independent invasion. This result argues that invasion genes contain sequence motifs that suffice to recruit Ste12 variants in the absence of the required cofactor Tec1. Because small changes within the Ste12 DNA binding domain shifted trait preference, we hypothesized that these changes allow Ste12 alone to occupy degenerate sites that would normally also require the binding of Tec1.

To explore this hypothesis, we cloned two canonical mating pheromone response elements (PREs, TGAAAC), derived from *STE12*'s own regulatory sequence, upstream of a minimal *CYC1* promoter driving expression of firefly luciferase. Addition of the PREs to this reporter gene increased luciferase expression 5-fold in the presence of the wild-type Ste12, but only 2.5-fold in the presence of the hyper-invasive variant SKQR (Fig. 3A). The failure of SKQR to fully activate may be due to its decreased affinity or ability to dimerize. We designed a second reporter gene containing a fully symmetric binding site (PRE^{sym}). Wild-type Ste12 activated the PRE^{sym} reporter to a far greater extent than it did the PRE reporter; in contrast, the SKQR variant performed even worse with PRE^{sym} than with PRE (Fig. 3A). These results demonstrate that Ste12 variants with opposing preferences for mating *(e.g.* wild-type) or invasion (*e.g.* SKQR) show differences in DNA-binding.

To analyze the DNA-binding preferences of Ste12 variants in more detail, we tested the ability of a library of ~12,000 binding sites, with random bases in the last position of the core and in two

3' flanking positions of each PRE, to mediate Ste12-dependent transcription of a reporter gene needed for growth under selection conditions (Fig. 3B). We carried out a library selection with four Ste12 variants: wild-type; SKQR, which is mating-proficient and hyper-invasive; K149E, which eliminates mating but is neutral for invasion; and K152L, which eliminates mating and is hyper-invasive. Wild-type and K149E preferred two copies of the canonical TGAAAC(A) to activate the reporter. In contrast, SKQR had a weaker preference for C in the last core position, while K152L showed a weak preference for A over C at this position (Fig. 3C); these two variants showed no preference at the first 3' flanking base (Fig. 3C). To determine the extent of Ste12-dependent activation due to these binding preferences, we grouped all sequences from the library containing either C or A in the last core position. Wild-type Ste12 and the mating-deficient K149E variant activated expression similarly from a dimeric TGAAAC, as is found in many genes required for mating (Fig. 3D, left). This surprising result indicates that the ability of Ste12 to bind to and drive expression on a binding site with two canonical PREs does not suffice for mating; rather, K149E may affect the interaction of Ste12 with a cofactor. In contrast, the hyper-invasive variants SKQR and K152L showed greatly decreased or no activation, respectively, at binding sites with canonical PREs (Fig. 3D, left), but they activated expression at sites with two copies of TGAAAA (Fig. 3D, right). Thus, the mechanism by which Ste12 variants increase invasion at the cost of mating may be their change in DNA-binding specificity to include TGAAAA motifs. We analyzed previous Ste12 chromatin immunoprecipitation data (*9*) and found a reduced preference for a C at core position 6 in genes bound during invasion relative to those bound during mating (Fig. 3 – Suppl. 2). Furthermore, invasion genes bound by Ste12 but not Tec1 show a similar decreased preference for C and a slight preference for A at this position (Fig. 3 – Suppl. 2). These data support the idea that in wild-type cells, Ste12 may adopt alternative binding modes, favoring binding sites that do not contain the canonical C found in the PRE.

Fungal pathogens like *Cryptococcus neoformans* transition toward an invasive lifestyle upon sensing the increased body temperature of their animal hosts (*13*). Having established that trait preference in *S. cerevisiae* can be modulated by genetic changes in Ste12, we asked whether temperature affected specificity toward mating and invasion. Furthermore, since chaperones maintain protein function at increased temperatures and modulate pheromone signaling in *S. cerevisiae* (*2*), we also asked whether mating or invasion changed in the presence of radicicol, a pharmacological inhibitor of Hsp90. Although the MAPK pathway contains the Hsp90 client Ste11, Ste12 is not a known client of this chaperone (*14*). We confirmed that mating is modulated by temperature and by Hsp90 function by mating cells expressing wild-type Ste12 in the presence of radicicol, or at increased temperature (Fig. 4A). We then subjected the Ste12 variant library to mating selection at 37°C or with Hsp90 inhibited. Most Ste12 variants responded to increased temperature or Hsp90 inhibition as wild-type Ste12 did (Fig. 4 – Suppl. 1). However, there were two positions, K150 and K152, in which mutations were highly temperature-responsive and Hsp90-dependent (Fig. 4B). This pair of lysines resides within the mutagenized segment that modulates mating and invasion specificity. We validated the temperature and Hsp90 effect on a variant (K150I) that mates near wild-type levels in the absence of heat or radicicol treatment (Fig. 4C). However, cells with the K150I variant under either treatment were severely decreased in mating. Thus, while Ste12 variants that are Hsp90-dependent are rare, they are accessible through a single amino acid change, suggesting that the chaperone could facilitate a mutational path toward the pathogenic lifestyle by minimizing

mating costs at 30°C and enhancing invasion at 37°C. Because K150 resides in the segment of the DNA-binding domain that modulates trait preference, we expected that invasion at high temperature would increase for variants at this position. We therefore conducted selection for invasion at 37°C. Indeed, the mean effect of all variants at K150 was to increase invasion at high temperature with a concomitant decrease in mating (Fig. 4D, 4E), demonstrating that K150 variants were not simply unstable at high temperature. Thus, mutations in Ste12 can interact with an environmental factor to further bias cellular decision-making toward invasion over mating.

Discussion

In summary, we carried out mutational scans for two independent traits to reveal specificity mediated through the DNA-binding domain of an ancient, fungal-specific transcription factor. This DNA-binding domain shows similarity to a homeodomain, but rather than a typical homeodomain recognition helix, it has a larger region predicted to contain two helices interrupted by a loop centered around the conserved W156 (*15*). The arrangement of mating- and invasion-preferring mutations on either side of the apparent recognition helix suggests that this DNA-binding domain offers flexibility to accommodate alternative binding modes to shift trait preference. The potential for mutations to generate alternative DNA-binding modes is supported by a recent study using protein binding microarrays to evaluate the specificity of disease-associated transcription factor variants (*16*, *17*). Rare coding variation in the homeodomain recognition helices of several transcription factors alters DNA binding specificity, which likely contributes to associated diseases in humans. Strikingly, some homeodomain variants bind at sites that differ only subtly from the wildtype, yet this promiscuity is associated with specific diseases (*17*). In other cases, the binding preferences of homeodomain and other transcription factors become subtly altered through formation of transcription factor pairs (*18*).

We demonstrate here that Ste12 can explore alternative binding modes through mutation, and thereby shift trait preference. However, throughout fungal evolution, the DNA-binding domain remains largely invariant, reinforcing the idea that the protein more typically takes advantage of both multiple interacting partners that have been gained and lost over time (*19*) as well as environmental conditions such as temperature to influence decision-making. Two of the most widespread fungal pathogens, *Candida albicans* and *Cryptococcus neoformans,* have adopted lifestyles that favor invasion of host tissues at the cost of mating (*20*, *21*). This tradeoff can be modulated by environmental conditions that mimic the body temperature of the host, with invasion favored when the pathogen shifts from 30^{o}C to 37^{o}C (*13*, *22*). That mutations affect trait preference as well as environmental sensitivity suggests that Ste12 may play a role in allowing fungi to rapidly adopt a pathogenic lifestyle.

Another common mechanism to generate trait specificity is gene duplication followed by subfunctionalization (*23*, *24*). Ste12 is very rarely duplicated in fungal genomes, which is consistent with our finding that hyper-invasive Ste12 variants are dominant over wildtype. Thus, we suggest that Ste12's atypical homeodomain reflects a requirement to accommodate both mating and invasion under the control of a single factor.

Materials and Methods

Generation of *STE12* mutant libraries. The *STE12* locus from *Saccharomyces cerevisiae* strain BY4741, including the intergenic regions, was introduced into the yeast vector pRS415 containing a *LEU2* marker (*25*). Degenerate DNA sequence encoding a 33 amino acid (99bp) segment of Ste12's DNA binding domain was generated by 2.5% doped oligonucleotide synthesis (Trilink Biotechnologies, San Diego, CA). Invariant 30bp sequences were designed on either side of the mutagenized fragment; these flanking sequences contained NotI and ApaI cut sites found in the coding sequence of *STE12*, and were unique in the *STE12* plasmid construct. Both fragment and plasmid were double-digested with NotI and ApaI, and the plasmid library was assembled by standard ligation. *STE12* libraries were transformed into electrocompetent *E. coli* (ElectroMAX DH10B, Invitrogen), and amplified overnight in selective media. Efficiency of ligation was verified by Sanger sequencing across the mutagenized region of 96 transformants (*26*); no assembly errors were detected, and mutant proportions reflected those expected for doped oligo synthesis at 2.5%. Plasmid libraries were used to transform yeast (BY4741 MATa or Σ1278b-α) with a deleted endogenous copy of *STE12* by high-efficiency lithium acetate transformation (*27*). The same plasmid library was used to transform yeast (Σ1278b-α) for invasion selection. Individual point mutations were generated in wild-type *STE12* plasmids by site-directed mutagenesis (Q5, New England Biolabs).

Large-scale trait selections. The BY4742 MATα strain was used as the mating partner for the library-transformed BY4741 MATa in all selections and mating assays. Transformed yeast cells were grown to late log-phase in a single 500mL culture, and cells were harvested to determine plasmid variant frequencies in the input population. The same culture was used to seed 36 independent mating selections for each treatment: one million MATa cells with *STE12* variants were mixed with 10-fold excess wild-type MATα cells and allowed 5 hours to mate(*28*). Depending on treatment type, cell mixtures were left at 30C with DMSO, 30°C with the Hsp90 inhibitor radicicol, or 37°C with DMSO. A 5uM concentration of the Hsp90 inhibitor radicicol (Sigma-Aldrich, R2146) was chosen due to its measureable effect on mating efficiency and lack of pleiotropic growth defects. Radicicol was chosen over the Hsp90 inhibitor geldanamycin (Sigma-Aldrich, G3381), because five-fold higher concentrations of geldanamycin were required to achieve the same phenotypic effect as with radicicol (data not shown). The temperature of 37°C was chosen for the similarity of effects on mating between temperature and radicicol treatments. After mating was completed, cell mixtures were plated using auxotrophic markers present only in mated diploids. Plasmids containing *STE12* variants were extracted from this output population, as well as from the pre-mating input, for subsequent deep sequencing. Mating selections were repeated in triplicate. For invasion, Σ1278b-α yeast cells transformed with the Ste12 plasmid library were grown to late log-phase in a single 500mL culture, diluted (10,000 cells per plate), and plated onto 40 plates of synthetic complete medium lacking leucine (2% agar). Plates were incubated at 30°C or 37°C for 72 hours to allow for sufficient invasion, as previously described (*29*). After three days, cells were washed from the plate surfaces, enriching for cells embedded in the agar. Agar pucks were removed from plates with a razor. Using the "salsa" blender setting (Hamilton Beach, Glen Allen, VA), a coarse slurry was generated and subsequently poured over a vacuum apparatus lined with cheesecloth. The resulting liquid cell suspension was spun down at 5000rpm to collect cells for subsequent deep sequencing. Individual invasion assays (as in Fig. 2F) were treated identically, but 10uL aliquots of OD-

normalized cultures were plated to image colonies for each strain. For high osmolarity growth, selections were conducted using the BY4741 MATa library-transformed population grown overnight in media containing 1.5M Sorbitol, as described previously (*30*). Populations were sequenced before and after growth to determine enrichment scores that defined the 95% confidence interval used in Fig. 1D. For all trait selections, we chose sample sizes that were at least 10-fold higher than the variant library size, ensuring that each variant would be adequately sampled in each of the three biological replicate selections.

Sequencing and determination of trait scores. Sequencing was completed on Illumina's MiSeq or NextSeq platforms. Sequencing libraries were prepared by extracting plasmids from yeast populations (Yeast Plasmid Miniprep II, Zymo Research, Irvine, CA) before and after selection. These plasmids were used as template for PCR amplification that added adaptor sequences and 8bp sample indexes to the 99bp mutagenized region for sequencing (all libraries amplified < 15 cycles). Paired-end reads spanning the mutagenized region were filtered to obtain a median of 5 million reads per sample. Using ENRICH software (*31*), read counts for each variant before and after selection were used to determine mating efficiency and invasion ability of *STE12* variants. Briefly, counts for a particular variant in the input and output libraries were normalized by their respective read totals to determine frequency in each, and a ratio of the output and input frequencies determine a variant's functional score. Finally, enrichment scores are normalized by the enrichment of wild-type Ste12 in each selection. Treatment scores are calculated identically, and in all cases where difference scores are shown, the score represents log2(treated) – log2(untreated).

Calculating intramolecular epistasis scores. Intramolecular epistasis scores were defined as the deviation of double-mutant's functional score (W_{ij}) from the multiplied scores of its constituent single-mutants (w_i* w_j). A negative epistasis score indicates that the deleterious effect of one mutation is increased by the presence of the partner mutation, while a positive epistasis score indicates that the partner mutation decreases the deleterious effects of individual mutations(*32*).

Quantitative mating assay. Individual variants tested for mating efficiency were treated identically to the large-scale mating selection experiments, except genotypes were scored individually on selective plates for either mated diploids (2N) or both diploids and unmated MATa haploids (2N, 1N). The proportion of mated individuals (mating efficiency) is taken as the ratio of colony counts on diploid to counts on diploid + haploid plates (2N/2N1N). Ste12 variants were always tested alongside wild-type to determine relative changes in response to treatment.

STE12 variant RNA-seq. RNA was extracted from yeast cells harboring the *STE12* mutant library grown under non selective conditions using acid phenol extraction as previously described(*33*). *STE12*-specific cDNA was created using a gene-specific cDNA primer and Superscript III (Life Technologies). cDNA was amplified in a manner similar to the plasmids, and prepared for sequencing using Illumina Nextseq.

Luciferase assays. BY4741 MATa *ste12*Δ yeast cells were transformed with two plasmids: one containing the *STE12* gene and one containing a reporter upstream of the firefly luciferase gene.

The promoter for the reporter contained two divergent PREs, derived from *STE12*'s own promoter, that are separated by three adenines, upstream of a minimal *CYC1* promoter. Although Ste12 does bind single PREs as a monomer, two sites are needed for signal detection in reporter assays (*34*). Yeast cells harboring both plasmids were grown overnight, back-diluted to identical optical density, and allowed 2-3 doublings before luciferase measurement. Luciferase activity was measured on a Biotek Synergy H1 plate reader immediately after addition of luciferin.

Large-scale analysis of Ste12 binding sites. A *HIS3* reporter gene was used for testing large populations of binding site variants (Fig. 3B) in media lacking histidine. We designed oligonucleotides that maintain the central portion of the native PRE, but randomized six surrounding bases on either side (NNNNNNTTTCAAAATGAAANNNNNN). This library was cloned into a promoter from the same plasmid used in the luciferase assay, at the same position as the native PRE. Strains containing one of four Ste12 protein variants were transformed with the same binding site library reporter population, and grown overnight in synthetic media lacking histidine with 10mM of the His3 competitive inhibitor 3-amino-triazole. Three biological replicate selections were conducted for each Ste12 protein variant tested against the binding site library. Cells were collected and sequenced before and after selections to determine enrichment scores for each binding site variant. Computational analysis of binding site enrichment scores was identical to pipeline for STE12 protein variants. Enrichment of all binding site variants are shown relative to the enrichments of empty plasmids, which were spiked-in to the binding site library as control. The 250 binding sites with the highest enrichments were grouped and Weblogo (*35*) was used to generate base preference plots.

Evolutionary analysis of Ste12 DNA-binding domain among fungi. Using fungal genomes deposited into NCBI and from the fungal sequencing project at Joint Genome Institute, we created a BLAST database of translated coding sequences and queried with the *S. cerevisiae* Ste12 protein sequence. Full protein sequences from BLAST hits were aligned using MUSCLE (*36*) and used to determine conservation at each site in the mutated segment. The secondary structure of *S. cerevisiae*'s Ste12 DNA-binding domain was determined using Psipred (*37*).

Analysis of Ste12 ChIP data. Genes bound by Ste12 only in mating conditions were compared to those bound by Ste12 only in invasion conditions (*9*). FIMO(*38*) was used to extract all matches to the Ste12 binding site in each of these gene sets. The frequency of the each possible base at position 6 of the core STE12 motif was determined relative to the genomic background (STE12 binding sites in all upstream sequences, no separation by mating or invasion gene function). The core 6 base frequencies relative to genomic background was then compared for genes bound in mating or invasion conditions. We used the same method to look at core 6 base frequencies at genes bound by both Ste12 and Tec1 or those bound by Ste12 alone during invasion.

References:

1. A. Mody, J. Weiner, S. Ramanathan, Modularity of MAP kinases allows deformation of their signalling pathways. **11** (2009), doi:10.1038/ncb1856.
2. J. F. Louvion, T. Abbas-Terki, D. Picard, Hsp90 is required for pheromone signaling in yeast. *Mol. Biol. Cell*. **9**, 3071–83 (1998).

3. J. G. Zalatan, S. M. Coyle, S. Rajan, S. S. Sidhu, W. a Lim, Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation. *Science*. **337**, 1218–22 (2012).
4. J. G. Cook, L. Bardwell, J. Thorner, Inhibitory and activating functions for MAPK Kss1 in the S. cerevisiae filamentous-growth signalling pathway. *Nature*. **390**, 85–88 (1997).
5. J. W. Dolan, C. Kirkman, S. Fields, The yeast STE12 protein binds to the DNA sequence mediating pheromone induction. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5703–7 (1989).
6. B. Errede, G. Ammerer, STE12, a protein involved in cell-type-specific transcription and signal transduction in yeast, is part of protein-DNA complexes. *Genes Dev.* **3**, 1349–1361 (1989).
7. Y. L. O. Yuan, I. L. Stroke, S. Fields, Coupling of cell identity to signal response in yeast: Interaction between the alphal and STE12 proteins. *Genes Dev.* **7**, 1584–1597 (1993).
8. H. D. Madhani, G. R. Fink, The control of filamentous differentiation and virulence in fungi. **8924**, 348–353 (1998).
9. J. Zeitlinger *et al.*, Program-Specific Distribution of a Transcription Factor Dependent on Partner Transcription Factor and MAPK Signaling. **113**, 395–404 (2003).
10. S. Chou, S. Lane, H. Liu, Regulation of Mating and Filamentation Genes by Two Distinct Ste12 Complexes in Saccharomyces cerevisiae †. **26**, 4794–4805 (2006).
11. O. Shoval, Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space. *Science (80-.).* **1157** (2012), doi:10.1126/science.1217405.
12. J. A. Fraser *et al.*, Same-sex mating and the origin of the Vancouver Island Cryptococcus gattii outbreak. *Nature*. **437**, 1360–1364 (2005).
13. J. R. Perfect, Cryptococcus neoformans : the yeast that likes it hot (2006), doi:10.1111/j.1567-1364.2006.00051.x.
14. R. Zhao *et al.*, Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell*. **120**, 715–27 (2005).
15. Y. L. Yuan, S. Fields, Properties of the DNA-binding domain of the Saccharomyces cerevisiae STE12 protein. *Mol. Cell. Biol.* **11**, 5910–8 (1991).
16. A. N. McKeown *et al.*, Evolution of DNA specificity in a transcription factor family produced a new gene regulatory module. *Cell*. **159**, 58–68 (2014).
17. L. A. Barrera *et al.*, Survey of variation in human transcription factors reveals prevalent DNA binding changes. *Science (80-.).* **351**, 1450–1454 (2016).
18. A. Jolma *et al.*, DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature*. **527**, 384–388 (2015).
19. T. R. Sorrells, L. N. Booth, B. B. Tuch, A. D. Johnson, Intersecting transcription networks constrain gene regulatory evolution. *Nature*. **523**, 361–365 (2015).
20. a Johnson, The biology of mating in Candida albicans. *Nat.Rev.Microbiol.* **1**, 106–116 (2003).
21. H. Dong, W. Courchesne, A novel quantitative mating assay for the fungal pathogen Cryptococcus neoformans provides insight into signalling pathways responding to nutrients and temperature. *Microbiology*. **144 (Pt 6**, 1691–7 (1998).
22. R. S. Shapiro, L. E. Cowen, Thermal Control of Microbial Development and Virulence : Molecular Mechanisms of Microbial Temperature Sensing. **3**, 1–6 (2012).
23. J. T. Bridgham, S. M. Carroll, J. W. Thornton, Evolution of hormone-receptor complexity by molecular exploitation. *Science*. **312**, 97–101 (2006).
24. M. Lynch, J. S. Conery, The evolutionary fate and consequences of duplicate genes.

Science. **290**, 1151–5 (2000).
25. D. Mumberg, R. Müller, M. Funk, Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*. **156**, 119–22 (1995).
26. F. Sanger, S. Nicklen, a R. Coulson, DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–7 (1977).
27. R. D. Gietz, R. a Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**, 87–96 (2002).
28. J.-Y. Leu, A. W. Murray, Experimental evolution of mating discrimination in budding yeast. *Curr. Biol.* **16**, 280–6 (2006).
29. O. Ryan *et al.*, Global gene deletion analysis exploring yeast filamentous growth. *Science*. **337**, 1353–6 (2012).
30. F. Posas, E. a Witten, H. Saito, Requirement of STE50 for osmostress-induced activation of the STE11 mitogen-activated protein kinase kinase kinase in the high-osmolarity glycerol response pathway. *Mol. Cell. Biol.* **18**, 5788–5796 (1998).
31. A. F. Rubin *et al.*, Enrich2 : a statistical framework for analyzing deep mutational scanning data (2016).
32. C. L. Araya *et al.*, A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16858–63 (2012).
33. J. T. Cuperus, R. S. Lo, L. Shumaker, J. Proctor, S. Fields, A tetO Toolkit To Alter Expression of Genes in Saccharomyces cerevisiae (2015), doi:10.1021/sb500363y.
34. T. Su, E. Tamarkina, I. Sadowski, Organizational constraints on Ste12 cis -elements for a pheromone response in Saccharomyces cerevisiae. **277**, 3235–3248 (2010).
35. G. Crooks, G. Hon, J. Chandonia, S. Brenner, NCBI GenBank FTP Site\nWebLogo: a sequence logo generator. *Genome Res*. **14**, 1188–1190 (2004).
36. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
37. L. J. Mcguffin, K. Bryson, D. T. Jones, The PSIPRED protein structure prediction server. **16**, 404–405 (2000).
38. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO : scanning for occurrences of a given motif. **27**, 1017–1018 (2011).

Acknowledgments: We thank J. Thomas for advice on analysis of fungal variation. We thank D. Fowler and H. Malik for comments on the manuscript. The work was supported by NIH grant R01GM114166 to C.Q. and S.F. M.W.D. was supported by an NSF Graduate Research Fellowship and a WRF-Hall fellowship. S.F. is an investigator of the Howard Hughes Medical Institute, which supported J.T.C.

Author Contributions: M.W.D., J.T.C., S.F., and C.Q. conceived and interpreted experiments. M.W.D., S.F., and C.Q. wrote the manuscript, J.T.C. and J.A.C. provided comments. M.W.D. conducted experiments and data analysis; J.A.C. contributed to high-throughput binding assay.

Data Availability. High-throughput sequencing reads have been submitted to NCBI SRA, awaiting accession number. Datasets with calculated enrichment scores for each tested variant in each condition, along with positional mean scores, and scores from binding site library selections are provided in source data files.

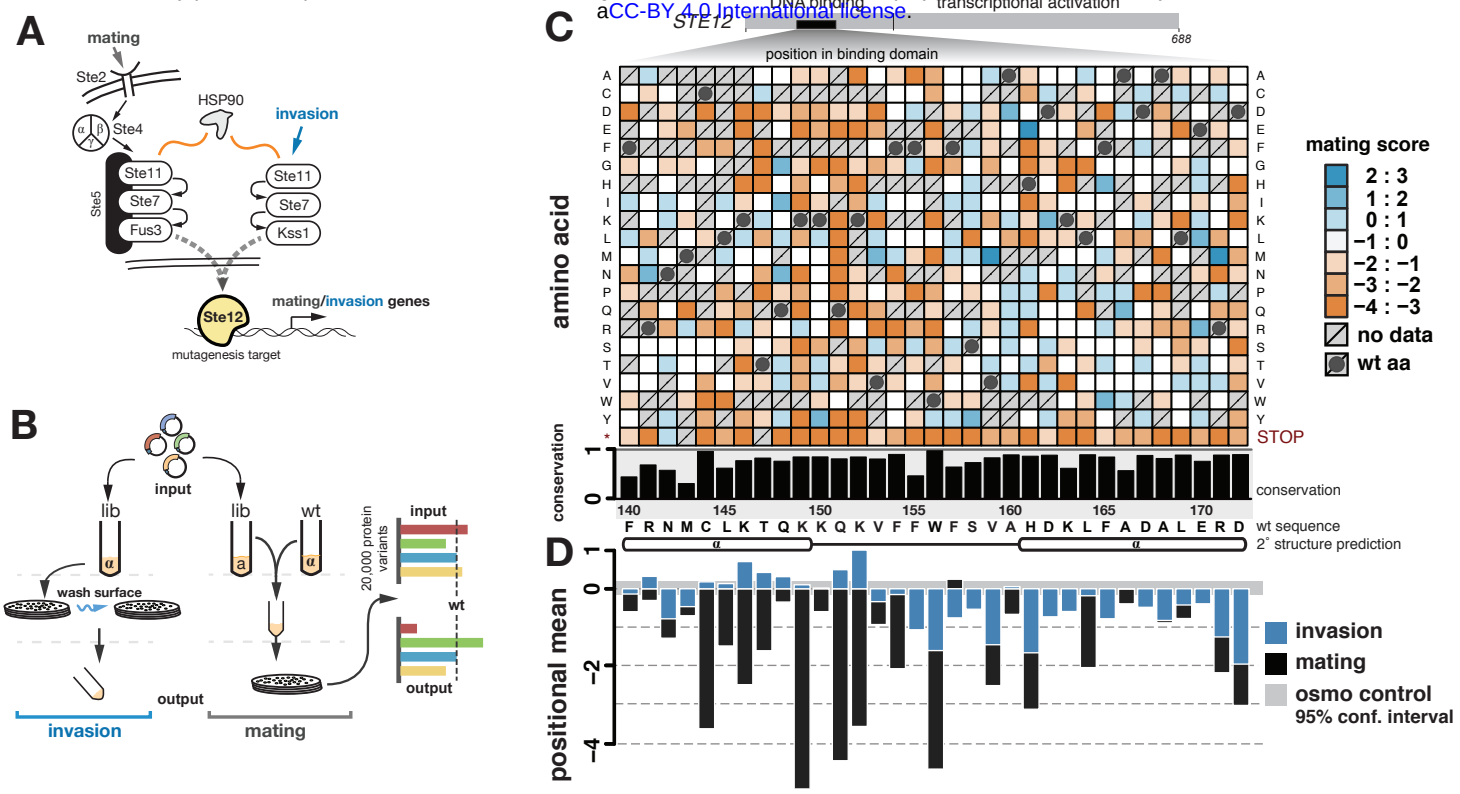

Fig. 1. Deep mutational scanning determines Ste12 DNA-binding domain function for both mating and invasion. (**A)** The yeast mating and invasion pathways contain shared signaling components, and both depend on Ste12 for activation of distinct regulatory programs. (**B**),Two yeast populations were transformed with the same STE12 variant library. For assaying invasion, SIGMA1278b MATα cells (n = 400,000 per replicate, 3 biological replicates) were plated on selective media, and grown. Selection was performed by washing colonies from plate surfaces and collecting cells embedded in the agar for sequencing. For assaying mating, BY4741 MATa ste12Δ cells were mated to MATα cells, and diploids (n = 500,000 per replicate, 3 biological replicates) were selected, scraped from the agar, and sequenced. In both cases, input variant frequencies were defined prior to selection. (**C)** The effects of single amino-acid substitutions within Ste12's DNA-binding domain on mating ability are shown. On the x-axis, the wild-type Ste12 sequence is shown, along with its predicted secondary structure (helices shown as tubes) and conservation. Conservation was determined as fraction of identity among 985 fungal species (Fig. S4). On the y-axis, amino acid substitutions are shown. Variants increasing mating efficiency are in shades of blue, and variants decreasing mating efficiency are in shades of orange based on log2 enrichment scores relative to wild-type. Dark grey circles indicate the wild-type Ste12 residue, and crosses indicate missing data. Ste12 variants showed comparable expression levels (Fig. 1 - Suppl. 4). (**D)** Positional mean scores for single amino acid substitutions are shown for mating (black) and invasion (blue), excluding stop codons. Grey horizontal bar indicates confidence interval for experimental noise determined from selection for Ste12-independent high osmolarity growth.

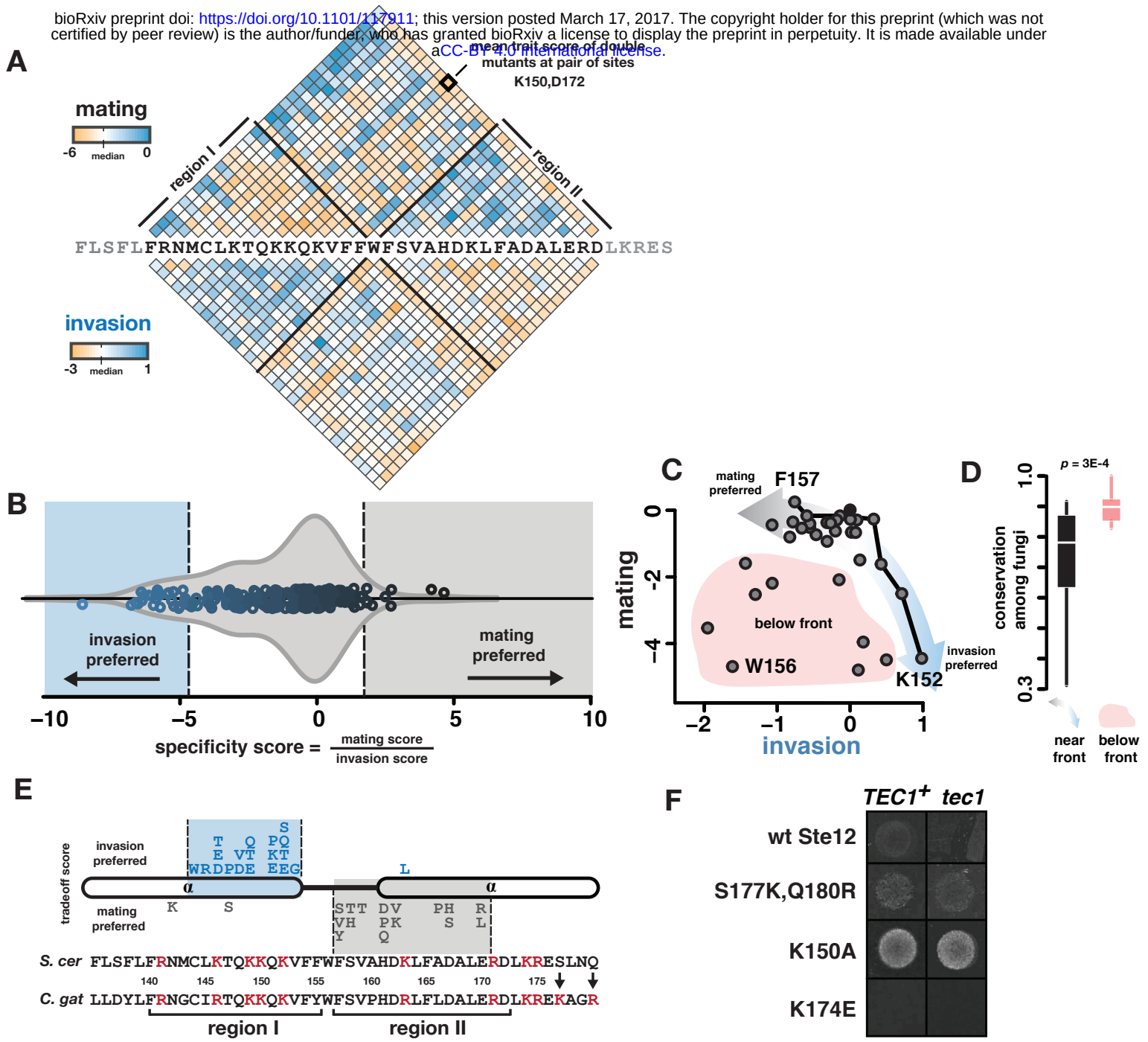

Fig. 2. Mutations in the Ste12 DNA-binding domain shift trait preference between mating and invasion. (A) Mean effect of double mutations between all combinations of positions for mating (above wild-type sequence) and invasion (below wild-type sequence). A specific pair of positions is shown in bold. Effects are color-coded as in Figure 1. Black lines emanating from W156 represent boundaries of region I, in which mutations primarily reduced mating, and region II, in which mutations primarily reduced invasion. (**B**) Along the x-axis, the specificity score for each mutation (circles) is superimposed on the full distribution (grey violin plot). The score is calculated as the log2 of the ratio of a mutation's effect on mating divided by its effect on invasion. S. cerevisiae Ste12 has a preference for mating: the largest density of mutations in the distribution have a negative specificity score, indicating that they increase invasion at the cost of mating (blue), while few mutations increase mating at the cost of invasion (black). Shaded boxes indicate mutations with the most extreme preferences (>2 standard deviations) for each trait. (**C**) Scatterplot of positional mean scores for both mating and invasion showed inverse relationship, indicating a tradeoff between both traits. The tradeoff is visualized as a Pareto front (black line), which was determined empirically; positions near the front maintain high values for one trait and minimize costs to the other. Arrows indicate preference for either mating (grey) or invasion (blue). Positions near the front were distinguished from those below the front (shaded in red) by calculating Euclidian distances. (**D**) Boxplots show conservation for positions near and below the front; positions below the front are more significantly more conserved among fungi than those near the front (two-sided t-test). (**E**) Single mutations from shaded boxes in B are mapped onto the predicted secondary structure of Ste12. Nearly every mutation that increased invasion at the cost of mating (blue) mapped to the first predicted helix, while most mutations that increased the preference for mating (black) mapped to the central loop and second predicted helix. S. cerevisiae and C. gattii sequences for this region are shown below, with positively charged residues in red. Arrows point to two additional positively charged residues found in C. gattii at the C-terminus of the region II helix that were used in the SKQR mutant. (**F**) Wild-type Ste12 requires the heterodimeric partner Tec1 for invasion, as a tec1Δ strain fails to invade. The region I mutation K150A led to invasion without Tec1. Introduction of the two positively charged residues found in C. gattii (SKQR) led to the same Tec1-independent phenotype, whereas introduction of a negative charge (K174E) eliminated invasion. Hyper-invasive mutants showed a dominant phenotype, as these mutants were tested in the presence of wild-type Ste12.

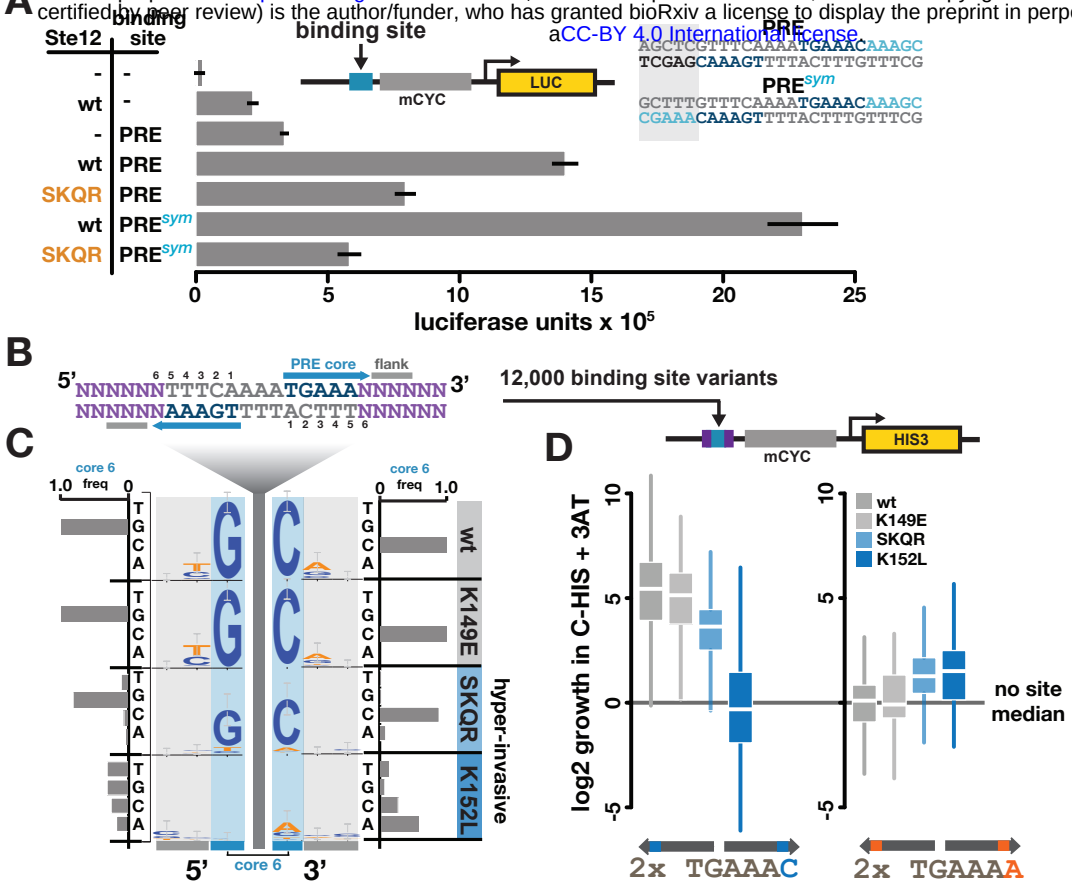

Fig. 3. Ste12 variants with increased invasion show altered binding-site specificity. **(A)** Wild-type Ste12 and the SKQR Ste12 variant were tested for activation of a reporter gene. The reporter contained either two PREs from the STE12 promoter or a version of this binding site mutated (boxed in grey) to perfect symmetry (PRE^{sym}). Core bases of the PREs are in dark blue and flanking bases in light blue. (**B**) A library of binding sites was cloned into a HIS3 reporter gene and used in growth selections. (**C**) The binding site library was subjected to selection in the presence of one of four Ste12 variants, wild-type and K149E (in grey), and two variants with increased invasion, SKQR and K152L (in blue). The sequence preferences for each Ste12 variant are shown as logo plots of three positions (core position 6 and two flansking bases, no preferences were seen at further flanking bases) of each PRE. Grey error bars on logo plots represent 95% confidence interval at each base. Bar charts show the frequency of base preference at core position 6. (**D**) In the left panel, the activity of each of the four Ste12 variants with all binding sites containing two canonical PREs is shown as a boxplot (n = 725, 747, 720, 608, respectively). In the right panel boxplots, the binding sites contained A instead of C at position 6 in the core (n = 696, 732, 775, 697, respectively). The altered preference of SKQR was validated with a quantitative growth assay (Fig. 3 - Suppl. 1).

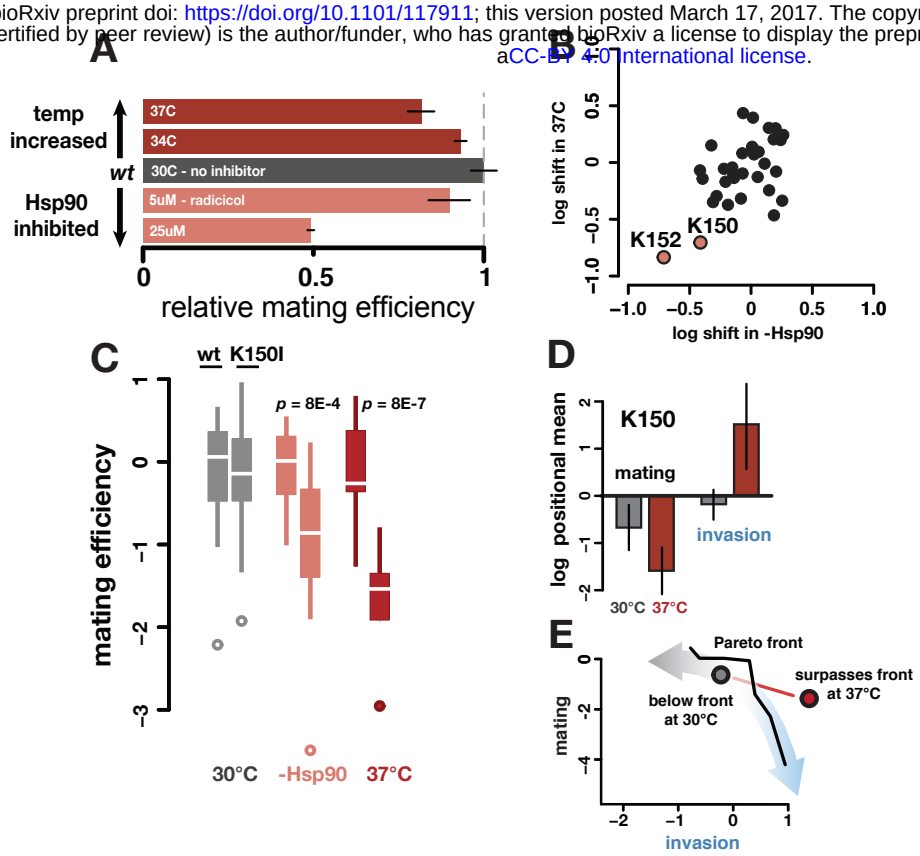

Fig. 4. Mutations in Ste12's DNA-binding domain enhance temperature-dependent modulation of trait preference. (**A**) Mating with wild type Ste12 at increased temperature (dark red) or in the presence of Hsp90 inhibitor radicicol (pale red) is reduced relative to an untreated control (black), error bars represent standard error of the mean (s.e.m). (**B**) The mating efficiency of Ste12 variants at high temperature or with Hsp90 inhibition is shown as the shift in mean effect at each mutated position. Two positions, K150 and K152 (pale red), showed greatest sensitivity to both treatments. (**C**) K150I was tested in a quantitative assay to validate its temperature-sensitive and Hsp90-dependent mating activity, shown relative to wild-type (left boxplot in each pair) in each treatment (n = 20 for each sample). **(D)** Ste12 variants at K150 (mean effect) decreased mating (left panel) and invasion (right panel) at standard temperature (grey bars). At high temperature (red bars), mating further decreased; however, invasion increased. Error bars represent s.e.m. The same data, represented as a scatterplot (**E**), show that K150, at high temperature, surpasses the Pareto front (Fig. 2C), increasing invasion at 37ºC with little cost to mating at 30ºC.

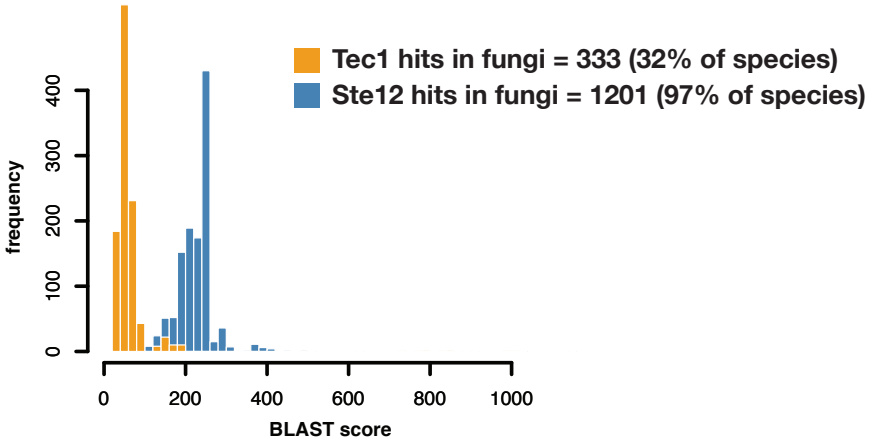

Fig. 1 - Suppl Fig 1. Tec1 is frequently lost in fungi whereas Ste12 is maintained. BLAST queries of 985 fungal genomes identified matches to either S. cerevisiae Tec1 (orange) or Ste12 (blue) protein sequence. Using a confidence threshold of BLAST score > 60, many fewer matches were found for Tec1, while most species maintained Ste12.

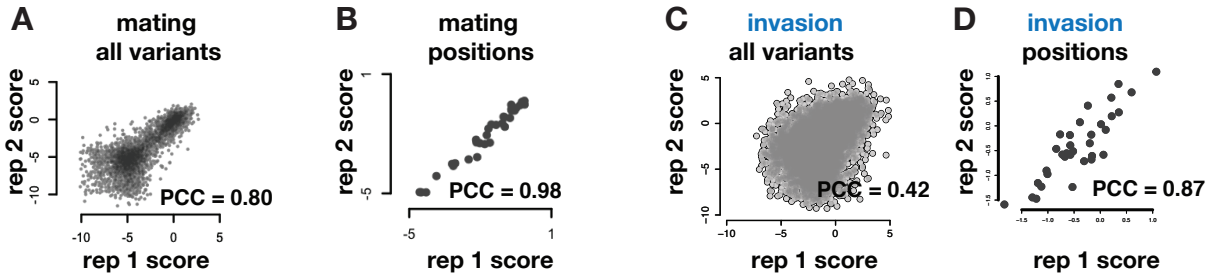

Fig. 1 - Suppl Fig 2. Mating and invasion enrichment scores are reproducible across replicate experiments. Selections for each trait were done in three biological replicates. We computed Pearson correlation coefficients (PCC) between replicates for all variants (**A,C**), as well as positional means (**B, D**). Shown are correlations for replicate 1 and 2, results were similar for other replicate pairs. Invasion selections were less well correlated. Nevertheless, positional mean scores were strongly correlated for both traits across replicates.

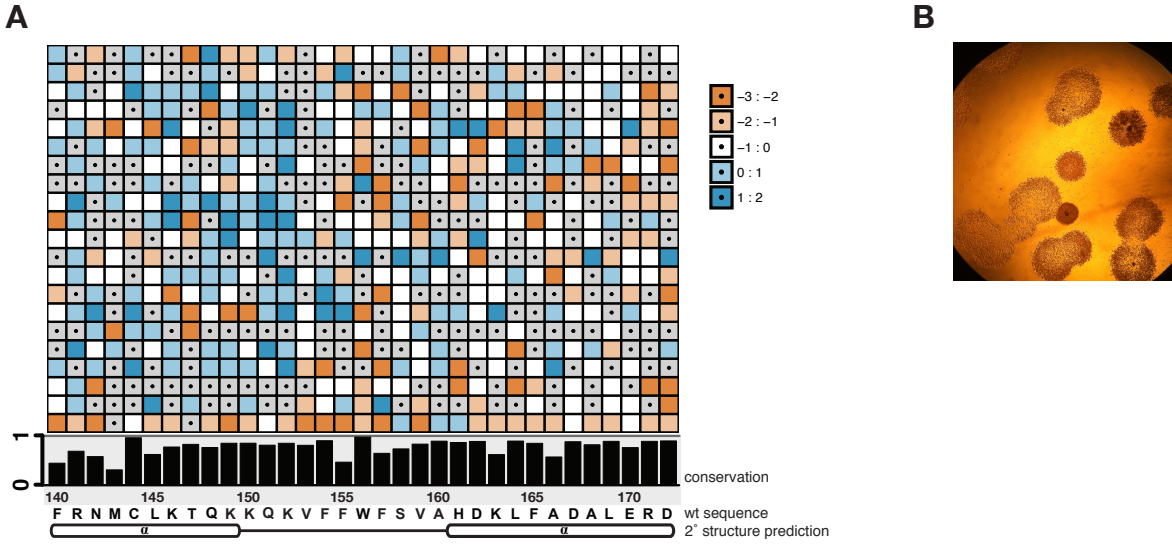

Fig. 1 - Suppl Fig 3. Variation in the highly conserved DNA binding domain of Ste12 generates a range of invasion phenotypes. (A) A heatmap of invasion scores for all single mutants is displayed as in Fig. 1C for mating scores. **(B)** A 10x magnified image of a washed plate containing invaded colonies with different Ste12 variants shows variation in the invasion phenotype.

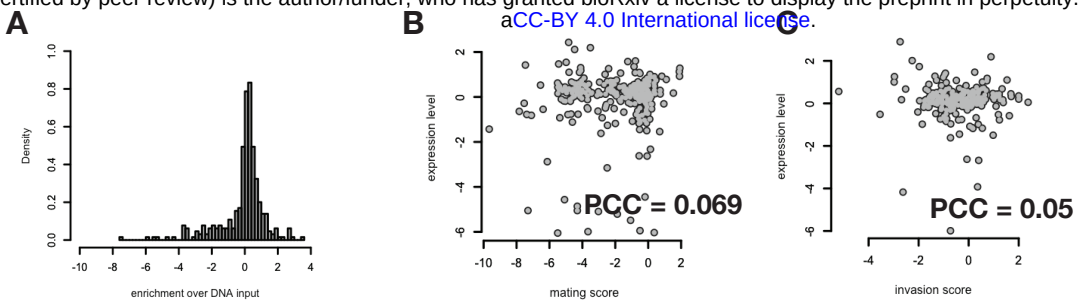

Fig. 1 - Suppl. 4. Variant expression shows little variation and no correlation with phenotypic effects. (**A**) The distribution of log2 enrichment of transcripts relative to plasmid counts for each Ste12 variant rarely deviates from zero, indicating low variation in expression level among variants. Each variant's expression level is plotted against each variant's trait enrichment score, for (**B**) mating and (**C**) invasion, demonstrating that variant expression levels did not affect phenotype.

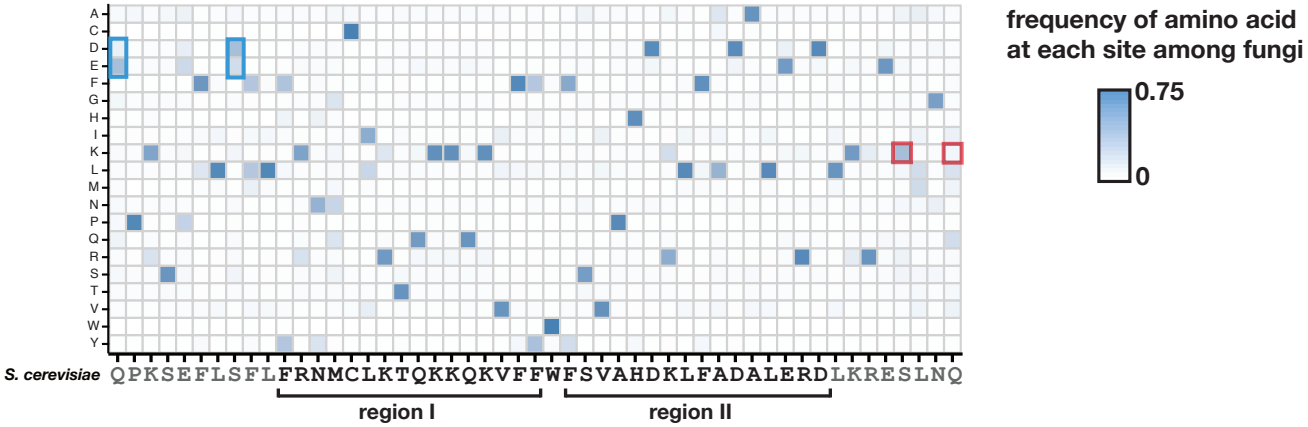

Fig. 2 - Suppl. 1. Ste12's DNA binding domain is highly conserved among fungi. We used a set of 985 fungal Ste12 protein sequences to examine the frequency of single amino acid changes in nature. Regions I and II defined in Fig. 2 are indicated. Note that S177 and Q180, positions C-terminal of region II that were altered in the hyper-invasive SKQR variant. Positively charged amino acids (red boxes) frequently occur at S177 and very rarely occur at Q180 compared to wild-type S. cerevisiae Ste12. Natural variation frequently introduces negatively charged residues (blue boxes) N-terminal of region I, in which mutations decrease mating and increase invasion. These observations support the assertion that wild-type S. cerevisiae Ste12 may promote mating more readily than invasion (Fig. 2B).

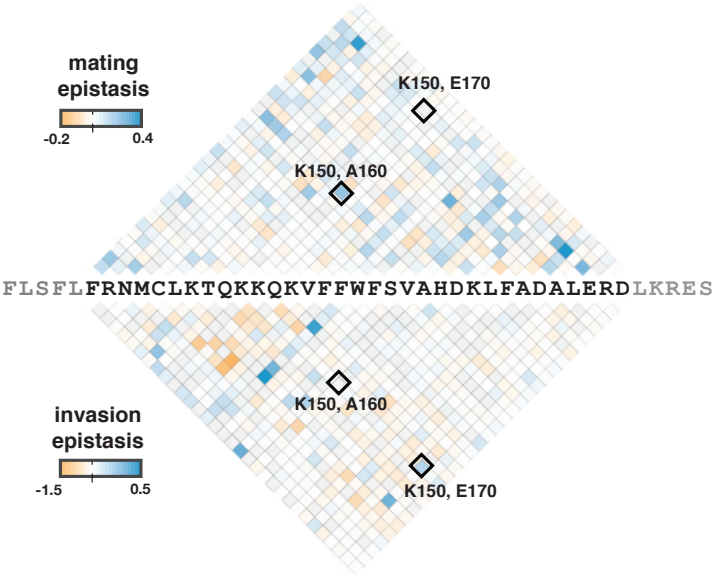

Fig. 2 - Suppl. 2. Double mutant analysis identifies pairs of positions with strong epistasis, and such pairs differ between mating and invasion. For each double mutant, we calculated epistasis as the deviation of that mutant's effect from the multiplied effect of its constituent single mutants (see Methods). Scores are displayed as in Fig. 2A; epistasis scores across all combinations of mutations tested at each pair of sites was used to calculate the displayed mean epistasis score. Sites with strong positive (shades of blue) or negative epitasis scores (shades of orange) differ between traits. Strong intramolecular epistasis occurs more frequently between residues in close proximity, suggesting that Ste12 conformation differs between both traits. Two pairs with trait-specific epistatic interactions are indicated in boxes.

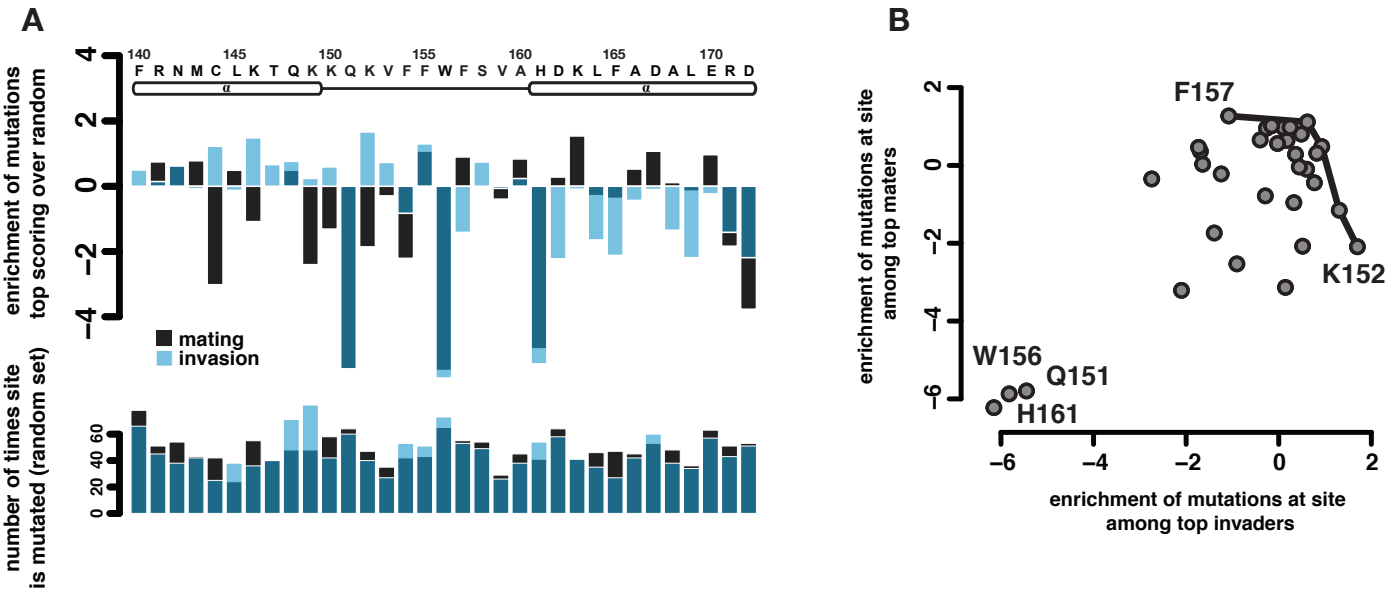

Fig. 2 - Suppl. 3. Orthogonal analysis reveals positions contributing to 'separation-of-function' and tradeoff between mating and invasion. Per-site mutation frequencies were calculated for the top 750 mating variants, top 750 invasion variants, and a random set of 750 variants from each dataset. (**A)** Enrichment of per-site mutations for top performing variants in each trait recapitulates bipartite arrangement of effects ('separation-of-function'), as well as sites increasing one trait at the cost of the other (e.g. see K152 and F157, tradeoff indicated with blue/black bars). (**B)** Per-site enrichment scores identify similar sites on the empirical Pareto front (Fig. 2C), as well as the same sites in which mutations are deleterious in both traits.

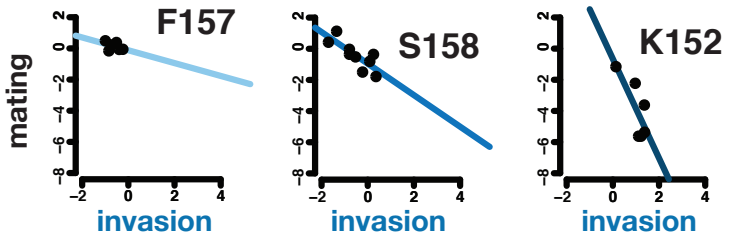

Fig. 2 - Suppl. 4. Orthogonal analysis reveals positions contributing to 'separation-of-function' and tradeoff between mating and invasion. Mating and invasion scores for all single amino acid change mutations at each indicated position were plotted, and a simple linear model was fitted. Steeper slopes indicate that mutations at that position are more likely to shift trait preference toward invasion.

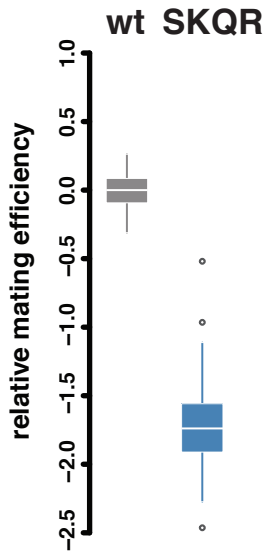

Fig. 2 - Suppl. 5. SKQR mutant shows decreased mating efficiency. Because the SKQR variant was not present in our initial Ste12 variant library, we used a standard quantitative mating assay to determine its mating efficiency relative to wild-type Ste12 as in Fig. 4C.

Fig. 3 - Suppl. 1. Validation of large-scale binding site selection scores. We tested different Ste12 binding sites with a range of activities identified in the large-scale sequencing-based assay in growth assays with a HIS3 reporter. (**A**) Growth rates of wild-type Ste12 with each binding site correlate with enrichment scores determined by sequencing. (**B**) The altered binding site preference of the SKQR variant is validated in individual growth assays. SKQR, but not wild-type Ste12, can activate a binding site with an A in core position 6.

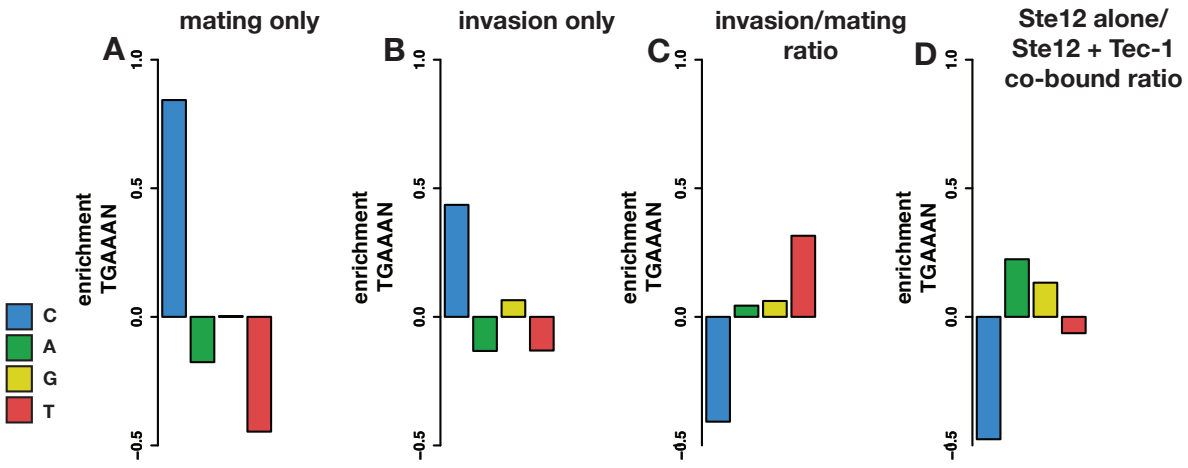

Fig. 3 - Suppl. 2. ChIP data analysis reveals altered binding site preference at yeast genes implicated in mating or invasion. TGAAA sites upstream of genes bound by Ste12 in mating (**A**) or invasion (**B**) were analyzed for the frequency of each base in core position 6. Base identities are color-coded. Enrichments were determined by comparing frequencies to those upstream of a random gene set. (**C**) Although both sets show enrichment for C in core position 6, invasion genes show reduced preference for C. (**D**) Repeating this analysis for invasion genes bound by either Ste12 alone or by both Ste12 and Tec1 also shows decreased C preference at Tec1-independent invasion genes.

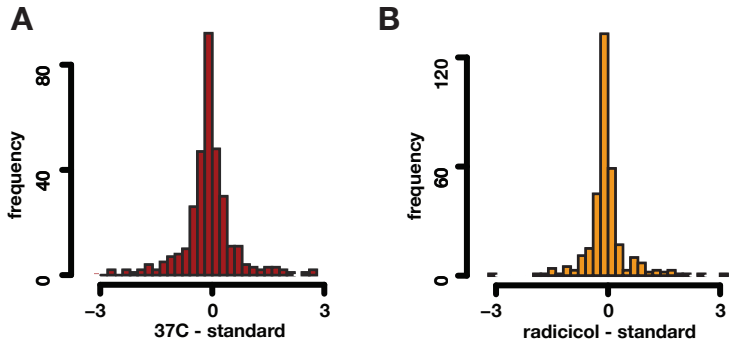

Fig. 4 - Suppl. 1. The majority of Ste12 variants respond like wild-type Ste12 to increased temperature or Hsp90 inhibition. The distribution of differences of log2 mating scores for all single mutants (treated - untreated) are plotted for high-temperature (red, **A**) and Hsp90-reduced (orange, **B**) conditions.