**Revealing Complex Genetic Bases of ABC Transporter Mediated Drug Resistance in Yeast using an Engineered Population Strategy**

**Abstract**

Unexpected effects of combined genetic perturbations (‘genetic interactions’) underlie complex traits and are critical to understanding genotype-phenotype relationships. ‘High-order’ genetic interactions­ - those involving three or more genes - ­are likely to greatly outnumber currently well-studied pairwise interactions, but are impractical to systematically investigate using current approaches. Here we demonstrate an ‘engineered population’ strategy to map arbitrarily complex genetic interactions in *S. cerevisiae* under multiple environments. We engineered and genotyped a population of ~7,000 uniquely DNA-barcoded strains with null variation segregating at 16 loci­ encompassing ATP Binding Cassette (ABC) transporters, and measured resistance of each to 16 drugs. We explored multi-knockout fitness effects and identified numerous high-order interactions conferring both drug resistance and sensitivity. Quadruple deletion of *SNQ2*, *YBT1*, *YCF1*, and *YOR1* was found to confer resistance to the *PDR5* substrates fluconazole and ketoconazole, pointing to a new model for complex mutual inhibition of ABC transporters, and new roles for *YBT1* and *YCF1*. We illustrate the potential for an engineered population strategy to systematically inform understanding of complex traits and gene function.

**Introduction**

A gene’s function in the context of a living organism is often understood by observing a resulting phenotype upon its perturbation. However, while some genetic perturbations result in straightforward and direct effects, many genotype to phenotype relationships are, in practice, a result of the combined influence of many genetic variants. An accurate understanding of a gene’s function, even in a controlled setting, must often involve discovering and accounting for potential unexpected effects when it is combined with other genetic variants. These effects can be broadly termed ‘genetic interactions’, and their presence implies that multiple simultaneous genetic perturbations can reveal phenotypic relationships which are absent when perturbing genes in isolation. Moreover, the presence of genetic interactions within a set of genes is often informative of some dependence in mechanism towards a biological function, such organization into pathways or complexes.

Because of the vast number of combinations, efforts to study ‘genetic interactions’ systematically, even in an organism with extensive availability of molecular tools such as as yeast, have been limited to two gene knockout combinations. These ‘two-gene’ interactions have been used to accurately cluster genes into functional groups1,2, better predict the genetic basis of drug resistance3, and to inform order-of-action in biological pathways 4. Understanding two-gene interactions may also be practically important for complex trait prediction in humans, as they may explain up to 80% of the missing heritability in some genome-wide association studies5.

Despite the success of two-gene interaction studies, many ‘high-order’ genetic interactions, those between three or more genes, have been discovered. It is expected that three-gene interactions (unexpected three-gene knockout effects) would greatly outnumber all discovered two-gene interactions if mapped genome-wide6. Perhaps reflecting this unexplored space and an extensive functional redundancy, approximately 1,000 genes in yeast show no two-gene interactions under standard conditions, and may require diverse environments and/or complex knockout combinations in order to understand their function2. Individual examples of high order interactions are known to arise from a variety of biological mechanisms, both simple and complex7, suggesting their broad applicability towards understanding genotype-phenotype relationships.

While the combinatorial complexity of high order interactions makes a genome-wide search impossible, they may be explored in-depth within a targeted set of genes. However, many current approaches to study genetic interactions involve individually engineering all combinations of desired genetic variants, which is inefficient for creating many knockout combinations involving multiple genes. Other approaches use controlled crosses of natural populations to introduce genome-wide variation *en masse*, which is then statistically associated to traits of interest3. However, such strategies are limited to associating naturally occurring variants of unknown causality, making biological interpretation and in-depth functional exploration limited. Furthermore, vast statistical power would be required to find complex interactions given a genome-wide scope.

Here we explore a method which combines strain engineering with a controlled cross to tractably map arbitrarily complex genetic interactions under multiple environments. Previously, we had deleted 16 ATP Binding Cassette (ABC) transporters in a single yeast strain (ABC-16)8. In this study, we have back-crossed the ABC-16 strain to a pool of isogenic wild-type ‘parents’, each containing a unique DNA barcode9. We genotyped 6,709 progeny from this cross *en masse* and linked this genotyping information to a strain-specific barcode, allowing us perform genotype-indexed growth profiling through high-throughput barcode sequencing under 16 drugs9. Overall, we performed population-level strain engineering within a targeted gene set in order to determine multi-knockout drug associations in what we term a depthwise association study (DWAS).

We discover many previously-unknown transporter-transporter relationships in the context of different drugs, even within this highly-studied gene family. Amongst our findings is a complex genetic basis of fluconazole resistance wherein a quadruple deletion of *snq2∆*, *ybt1∆*, *ycf1∆*, and *yor1∆* confered unexpected resistance which was dependent on *PDR5*. We confirm this complex genetic interaction through individual strain growth profiling, and perform several follow-up experiments guided by the genetic data to propose several possible underlying mechanisms.

**Results**

**Creation of a Genotyped and Phenotyped Engineered Population**

We first mated the ABC-16 strain to a pool of isogenic wild-type parents, each containing two DNA tags (‘barcodes’). After isolating approximately equal numbers MAT**a** and Mat**α** progeny from this cross, we performed large-scale genotyping using high-throughput sequencing of tagged PCR products, which was calibrated by comparison to 38 strains genotyped using individual PCR reactions (Methods; Fig. 1). We estimate a genotyping accuracy of 95% or higher for 12/16 genes, and an overall accuracy of 93.2% (Fig. S1A, Methods). Overall, we obtained useful genotyping data linked to at least one unique DNA barcode for 6,709 progeny with 5,095 unique genotypes, representing approximately 8% of the 216 (65,536) possible knockout combinations within this gene set.

We then tested the genotyping data for the presence of linkage between ABC transporters. We found that except for *BPT1* and *YBT1*, all genes were either unlinked or weakly linked (Fig. S1B). Three pairs of genes – *YOR1-YCF1*, *YOR1-BPT1*, and *SNQ2-PDR5* – further had weak, but (given the large sample size) significant apparent negative linkage, suggesting a mild synthetic growth defect under our no-drug growth condition (Fig. S1B).

We further assessed the distribution of knockouts amongst the progeny. The average number of knockouts in the pool was 7.0, which is most consistent with an estimated asymmetric genotyping accuracy of 94% (compared to 93% derived by comparison to gold standards; Fig S1C, Methods).

After linking each genotype to a strain-specific barcode, we pooled our arrayed collection in order to estimate competitive growth rates using high throughput barcode sequencing10 at five different time points (0, 5, 10, 15, 20 generations of overall pool growth, Fig. 1). The two pools were grown in the presence of 16 different drugs and a solvent control. Sequencing data from all time points were combined into an overall growth metric, which was used to compare the growth for each strain in each drug to the solvent control, yielding a resistance score (Methods, Fig. 1). After excluding all strains which were not present at t=0 in the solvent control, we obtained resistance scores for 3,186 strains in the MATa pool and 3,523 strains in the Matα pool. Here we treat these pools as independent population replicates. Further taking advantage of the fact that most strains had two successfully mapped unique DNA barcodes (‘UP’ and ‘DN’), we treated these as internal replicates within each pool. We found the resistance scores to be highly correlated (minimum r > 0.9 for MATa and MATα) amongst different barcodes for the same strain in both pools for 8 drugs, and adequately correlated (minimum r > 0.7) for 4 other drugs (Fig S2). This indicates adequate sequencing coverage, and that our resistance metric is not substantially affected by the specific DNA barcode sequence. For the remaining 4 drugs, the low correlation appeared to result from a lack of strain-strain variability in resistance rather than measurement noise, suggesting that the 16 ABC transporters tested do not play a major role in mediating resistance to the compound (Fig S2).

**Simple Drug Sensitivity to Knockout Associations in an Engineered Population**

As a preliminary test of using TWAS to associate gene knockouts to drug sensitivity, we first used our pools to recover marginal associations (i.e. association of a knockout in a population to drug resistance, aggregated over all genetic backgrounds) using a linear model. This ‘marginal TWAS’ recovered XX% of previously-known knockout associations, which had been limited to involving *snq2∆*, *pdr5∆*,and *yor1∆*. Previously-unreported associations with *ycf1∆*, *bpt1∆*, and *ybt1∆* were also common amongst the drugs tested, however their relativey small marginal effect sizes suggest that these genes have either minor knockout effects, and/or that their effects are dependent on the presence of other knockouts (XX Fig. SX XX). We found XX% of significant marginal associations to be reproducible between the MATa and MATα pools (XX Fig. SX XX), and further found that models trained on a pool of one mating type had comparable predictive power when applied to the opposite mating type for the same drug (XX Fig. SX XX). Thus, in a limited test scenario, TWAS recovered many previously-known associations, and suggests a role of three new genes towards resistance to the tested drugs.

**A Complex Fitness Landscape: Visualization of Surprising Multi-Gene Knockout Effects**

Given the marginal TWAS results, we used the population data to explore all multiple-knockout effects amongst *snq2∆*, *pdr5∆*, *yor1∆*, *ycf1∆*, *bpt1∆*, and *ybt1∆*. We grouped the strains in our pool based on all 26 knockout combinations amongst these transporters (allowing variation at the other 10 transporters within each group), in order to create a reduced population profile for each drug (XX Fig. SX XX). This profile was used to cluster drugs based on their multi-knockout sensitivity patterns (XX Fig. SX XX).

We then converted this grouped profile into a ‘complex fitness landscape’, which shows not only the average resistance of each knockout group, but also predicts the effects of knocking out any of the transporters in the context of any other knockout combination within the grouped genes. Even without providing an explicit model of genetic interactions, this landscape revealed many surprising reproducible genetic relationships in different drugs (XX Fig. ?? XX). In the most straightforward cases, multiple knockouts appeared to confer surprising drug sensitivity, which was not clearly present upon knocking out a smaller set of the underlying genes. Such effects are illustrated by the *snq2∆pdr5∆* group under camptothecin, and *snq2∆pdr5∆ybt1∆* under mitoxantrone (XX Fig. ?? XX). Broadly, such patterns suggest parallel action of these transporters in the efflux of their respective drugs.

In more complex cases, a knockout appeared to have opposite drug resistance effects depending on the genetic background. Under benomyl, for example, *yor1∆* was shown to have no significant effect by itself, to increase fitness relative to *pdr5∆* as *yor1∆pdr5∆*, and to decrease fitness relative to *snq2∆* as *yor1∆snq2∆* (XX Fig. ?? XX). While the first two results may be expected given a dependence of benomyl resistance to SNQ2 and an earlier report of a (non-additive) ‘compensatory activation’ of Snq2 given disruption of PDR5 and YOR111, the latter is a surprising and previously-unreported effect which suggests that Yor1 can serve a secondary role in benomyl resistance, for example by providing some efflux in the absence of Snq2.

In perhaps the most surprising examples, a complex knockout combination appeared to confer considerably improved drug resistance rather than sensitivity. Under fluconazole and ketoconazole, a quadruple deletion of *snq2∆ybt1∆ycf1∆yor1∆* led to drug resistance, whereas *pdr5∆snq2∆ybt1∆ycf1∆yor1∆* was comparably as sensitive as the *pdr5∆* group (XX Fig ?? XX)*.* Furthermore, combinations of one or two knockouts within these four genes appeared to have only minor effects (XX Fig ?? XX).

Given the potential of these grouped profiles to show unexpected multi-knockout effects, we developed two additional represenations. In the first representation, we created a ‘radial fitness landscape’ which maintains the knockout relationships present in a traditional fitness landscape representation, but in a compact visualization which allows an overview of all data and even quicker identification of multi-knockout effects (XX Fig ?? XX). We further observed the underlying distribution of drug resistance for 25 combinations of *snq2∆*, *pdr5∆*, *yor1∆*, *ycf1∆*, and *ybt1∆* under fluconazole(XX Fig ?? XX), further highlighting the surprising multi-knockout resistance combinations amongst these five genes. In many cases the underlying populations within each group do not appear to be normally-distributed, suggesting further context-dependent roles for the other 10 transporters in this drug which were not recovered by a test for marginal association.

**Formal Modelling of Complex Genetic Interactions: Complex Drug Sensitivity to Knockout Associations**

Given the observation of many apparently-surprising multi-knockout drug resistance phenotypes, we aimed to formally define and model complex genetic interactions within the engineered population. To do this, we developed a linear modelling approach, which extends the multiplicative fitness model of pairwise genetic interactions12 to incorporate interactions of arbitrary complexity (See Methods). We then treated the search for complex genetic interactions as a feature selection problem, and searched for up to to 4-gene interactions within a total space of 2,516 coefficients (See Methods; XX Fig ?? XX). We found XX 2-way, XX 3-way, and XX 4-way interactions which were reproducible between MAT**a** and MAT**α** pools at a stringent Bonferroni corrected p-value cutoff of 0.05 (XX Fig S? XX).

This approach formally captured the surprising genetic relationships found in the fitness landscape. For example, *yor1∆* was found to have no main effect under benomyl, to have a positive genetic interaction with *pdr5∆*, and to have a negative genetic interaction with *snq2∆* (XX Fig S? XX). In camptothecin, *pdr5∆* and *snq2∆* each had a minor main effect, as well as a strong negative interaction between them (XX Fig S? XX). Similarly, the surprising *pdr5∆snq2∆ybt1∆* phenotype in mitoxantrone was modelled as a combination of a small marginal effect of *snq2∆*, a negative interaction between *pdr5∆* and *snq2∆*, and a further triple negative interaction upon *YBT1* deletion as *pdr5∆snq2∆ybt1∆* (XX Fig S? XX). A similar ‘redundant transporter action’ phenotypic pattern was observed for *pdr5∆snq2∆yor1∆* in cisplatin (XX Fig S? XX). Under fluconazole, the resistance phenotype was modelled (in addition to one- and two- gene effects involving *pdr5∆*, *yor1∆*, *ybt1∆*, *pdr5∆yor1∆*) as the combination of three positive three-way interactions (*ybt1∆ycf1∆yor1∆*, *snq2∆ybt1∆yor1∆*, *snq2∆ycf1∆yor1∆*). The *PDR5* dependence of these effect were then further modelled as three additional negative four-way interactions(*pdr5∆ybt1∆ycf1∆yor1∆*, *pdr5∆snq2∆ybt1∆yor1∆*, *pdr5∆snq2∆ycf1∆yor1)*.

**Hypothesizing Drug-Dependent Transporter-Transporter Relationshipss from Multi-Knockout Phenotypes**

We sought to intepret all drug-dependent transporter-transporter relationships suggested by the multi-knockout data, using a combination of the grouped fitness landscapes and the formally-defined genetic interactions. The interpreted relationships were highly drug-dependent, and were derived from either two, three, or four knockout phenotypes (XX Fig ? XX, XX Data SX XX). All relationships involved either parallel clearance of a drug by two or more transporters, and/or the antagonism of an ABC transporter by the presence of another (XX Fig ? XX). Again, the genetic relationships further suggest an extensive role for YBT1 and YCF1 in many drugs, despite none having been previously characterized. YBT1 is modelled to work in parallel with SNQ2, PDR5, or YOR1 depending on the compound, and to antagonize PDR5 in the context of fluconazole resistance (XX Fig ? XX). YCF1 is modelled to antagonize SNQ2 and PDR5 in some compounds, and to work in parallel with SNQ2 in the context of benomyl resistance (XX Fig ? XX).

**Single-Strain Validation of Population-Based Genetic Interactions**

In order to validate the complex genetic interactions present in the population, we sought to generate all knockout combinations of *pdr5∆, snq2∆, yor1∆, ybt1∆,* and *ycf1∆* in individual strains*.* Through a targetted mating and genotyping strategy, all 32 knockout combinations were obtained (see Methods). We sought to use this collection to first confirm the complex interaction observed under fluconazole. We tested the resistance of each strain to fluconazole at various concentrations. Fluconazole resistance of these strains correlated well with the pooled resistance at (r=XX, Fig XX). As predicted by the genetic interaction in the TWAS pool, the *snq2∆yor1∆ybt1∆ycf1∆* strain was the fastest growing, and only minor resistance effects were seen for strains with one and two knockout combinations amongst these four genes. To test whether the lethality of the *pdr5∆snq2∆yor1∆ybt1∆ycf1∆* strain was a consequence of the quadruple deletion being unable to rescue the *pdr5∆* defect, we tested lower concentrations of fluconazole where *pdr5∆* strains still exhibited some growth. At these low concentrations (XXμm), all *pdr5∆snq2∆* strains, including *pdr5∆snq2∆yor1∆ybt1∆ycf1∆* grew worse than *pdr5∆* strains (Fig XX), showing that the resistance effect is dependent on PDR5, and suggesting that *snq2∆* can act as a minor efflux pump for fluconazole.

We used this pool to test other drugs used in the TWAS study. To contrast the

Additionally, we created an alternative map which models the effects of ‘adding in’ gene combinations to the ABC-16 strain, rather than the effect of ‘knocking out’ gene combinations from the wild-type strain (XX Fig S? XX). Within a redundant gene set such as these transporters, interactions from this model may make interpretation of some genetic effects more straightforward. For example, a positive effect of a gene add-in under this model may be interpreted as that gene having the ability to efflux a drug, even though its knockout from a wild-type strain may not manifest a discernable fitness defect because of the presence of other redundant transporters. ‘Add-in’ interactions between two transporters may then be further be interpreted in several ways depending on the direction of the main effects and the interaction – for example, the transporters may be working in parallel, or one transporter may inhibit or activate the other (see Methods).

In general, the found complex interactions had an effect size distribution with fewer ‘high-effect’ outliers, but a similar median effect size when compared to the single-gene effects and two-way interactions (XX Fig ?? XX).

**Confirming a Genetic Interaction of SNQ2, YBT1, YCF1,**

**Materials and Methods**

**Yeast Strains**

RY0622 (Green Monster MAT**a**):

*MAT***a** *adp1Δ snq2Δ ycf1Δ pdr15Δ yor1Δ vmr1Δ pdr11Δ nft1Δ bpt1Δ ybt1Δ ynr070wΔ yol075cΔ aus1Δ pdr5Δ pdr10Δ pdr12Δ can1Δ::GMToolkit*-**a** (*CMVpr-rtTA KANMX4 STE2pr-Sp-his5*) *his3Δ1 leu2Δ0 ura3Δ0 met15Δ0*.

RY0148 (Barcoder Strain MAT α):

*MAT α lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::GMToolkit-α (CMVpr-rtTA NatMX4 STE3pr-LEU2)*

**Media**

**DNA Oligomers Used**

LoxP:

Lox2272:

US1:

US2:

DS1:

DS2:

U1:  
U2:

D1:

D2:

5’HO:

3’HO:

**PCR Conditions**

For each PCR reaction performed, cells were lysed in a 96-well format in 2 μl of buffer (0.1 M sodium phosphate (pH 7.4) and 1 unit zymolyase (ZymoResearch) overlaid with mineral oil, incubating them at 37 °C for 20 min and then at 95 °C for 5 min.

**High Throughput Sequencing**

**Creating the Barcoder Plasmid**

A plasmid backbone expressing GAL1pr-CRE was combined with a ‘barcoder locus’ flanked by LoxP and Lox2272 sites using in-yeast-assembly. This barcoder locus consists of a random 25bp DNA sequence (‘UPTAG’) in between two common primer regions (US1 and US2), followed by a HphMX4 cassette, and then by another random 25bp DNA sequence (‘DNTAG’) in between two common primer regions (DS1 and DS2).

**Generating a Barcoder Strain**A URA3 cassette flanked by LoxP and Lox2272 sites was integrated into the HO gene of the RY0148 strain to serve as the locus for barcode integration. This strain was further transformed with the barcoder plasmid, and transformants were selected using HygroB containing media. These transformants were then subjected to a 4 day galactose induction to allow for Cre expression. All strains which had either lost the plasmid or were unsuccessful in Cre-mediated recombination were counter-selected by growth in the presence of 5-FOA. Strains surviving the counterselection were then grown in YPD in the presence of HygroB in order to select for successful integration of the plasmid barcode locus. Ten colonies were tested for barcode integration using two sets of primers to verify the strain barcode-specific UPTAG and DNTAG - US2 and a sequence complementary to 5’ of the HO gene (5’HO); DS1 and a sequence complementary to the 3’ of the HO gene (3’HO). PCR products were analyzed using gel electrophoresis (Supplementary Figure XX).

**Mating, Sporulation, and Haploid Selection**

For mating, we combined MAT**a** and MATα cell cultures 1:1 in <50 μl YPD at 1 OD600 nm, centrifuged the mixture at 735*g* for 5 min and incubated it at 30 °C for 24 h. We transferred 10 μl of the mating mixture to 500 μl GNA medium (starting OD600 nm ≅ 0.1) containing G418 and Nat, and cultured it for 24 h. This allowed for the selection of diploids owing to resistance to G418 and Nat conferred by *GMToolkit*-**a** (*KanMX4*) and *GMToolkit*-α (*NatMX4*), respectively. We transferred 20 μl of the 1-d culture to 500 μl of fresh GNA medium containing G418 and Nat (starting OD600 nm was ~0.2) and cultured for 5–7 h to bring cells to the log phase (OD600 nm was ~1) before sporulation. At this point, we rinsed cells three times with 500 μl of minimum sporulation medium (XX cite) and resuspended them in 1 ml of minimum sporulation medium containing 7.5 μg ml−1 lysine, 7.5 μg ml−1 leucine, 5 μg ml−1 histidine, 5 μg ml−1methionine, and 1.25 μg ml−1 uracil (to meet auxotrophic requirements). The sporulation mixture was rotated at room temperature (20–25 °C) for 1d and then at 30 °C for 3d (XX cite). We treated 125 μl of this mixture with zymolyase (2 units) at 30 °C in a 50-μl reaction (100 mM sodium phosphate buffer (pH 7.4) and 1 M sorbitol) for 1 h, followed by 5 min of treating with NP-40 (added to the zymolyase reaction to achieve a final concentration of 0.01%), and stopped the zymolyase reaction by adding 500 μl of water and placing it on ice. This mixture was sonicated at the output setting of 1 for one minute with the 50% duty cycle using Sonifier 450 (Branson), split equally into two tubes and centrifuged at 735*g* for 5 min.   To the pellet, we added 100 μl of SC without histidine (SC –His) or SC –Leu to separately select for haploid cells of MAT**a** or MATα, respectively, for 16 to 24 h, using the haploid selection markers (*STE2pr-Sp-his5* and *STE3pr-LEU2*).

**Generating Barcoded Random Knockout Progeny**

Mating, sporulation, and haploid selection were performed between the RY0622 ‘Green Monster’ strain (MAT**a**) and the barcoder strain (MATα). Using colony plating, sporulation efficiency was estimated at 24% - 1080 colonies grew in SC, 140 colonies grew in SC-His (MAT**a** haploid selection), and 120 colonies grew in SC-Leu (MATα haploid selection). The haploid pools were plated as a form of quality control. From the Sc-His pool, 440 colonies grew in Sc-His, 0 grew in Sc-Leu, 468 grew in YPD+G418, 13 (~2%) grew in YPD+Nat; from the Sc-Leu pool, 0 colonies grew in Sc-His, 1600 grew in Sc-Leu, 50 (~3%) grew in YPD+G418, 1650 grew in YPD+Nat. The two pools were then grown in YPD+HygroB to select for barcoded haploids. The Sc-Leu pool was further grown in Sc-Ura to select against barcoder strain parents that may have escaped diploid selection. Using a QPix robot, 5,376 colonies were picked from the MAT**a** pool and 6,144 colonies were picked from the MATα pool onto 384 well plates.

**Individual Strain Genotyping**

To genotype each strain at the 16 ABC transporter loci, two PCR reactions were performed for each locus - one to determine the presence of GFP cassette integration, and another to determine the presence of the wild type gene.

For the first set of reactions, locus–specific PCR primers from the 5′ flanking sequences of each gene were paired with a common primer complementary to the *GFP* cassette. For the second set of reactions, the 5’ locus-specific primers were paired with primers which were complementary to the internal sequence of their corresponding gene.

PCR products were analyzed using gel electrophoresis.

**Creating A ‘Gold Standard’ Genotyped Set**

To create a ‘Gold Standard’ genotyped set, 38 strains were randomly picked from the collection and individually genotyped. The strain-specific UPTAG and DNTAG were PCR amplified using two sets of primers – US2 and a sequence complementary to 5’ of the HO gene (5’HO); DS1 and a sequence complementary to the 3’ of the HO gene (3’HO). The PCR products were subjected to Sanger sequencing.

**Pooled Strain Genotyping**

We developed a strategy to perform *en-masse* genotyping of the random knockout progeny using high throughput sequencing.

In order to uniquely tag PCR products originating from the same well, all forward primers had a 5’ tag encoding the well row (R), and all reverse primers had a 5’ tag encoding the well column (C).

For each well in the collection, four PCR reactions were performed with the following primers: R+U1 and C+U2 to amplify DNA barcodes encoding the UPTAGs for each gene deletion; R+D1 and C+D2 to amplify the deletion-specific DNTAGs; R+US1 and C+US2 to amplify the strain-specific UPTAG; R+DS1 and C+DS2 to amplify the strain-specific DNTAG. PCR reactions were performed using a Hydrocycler. PCR products were pooled for each plate, and an additional PCR was performed with each plate pool in order to add a plate-specific Illumina sequencing adapter (P tag).

**Analysis of Pooled Strain Genotyping Data**

The pooled PCR products were sequenced using an Illumina MiSeq, and the reads were demultiplexed into individual samples corresponding to a plate and well of origin using a Perl script.

For each sample, we developed a pipeline to determine the strain-specific TAG sequences and genotype from the reads. The parameters of the genotype caller from this pipeline were trained using samples which matched the ‘Gold Standard’ set, and the cross-validated accuracy is reported in the text for each gene.

UPTAG or DNTAG identity and a corresponding genotype was successfully determined for 7,195 samples. For 7,030 samples, the UPTAG or DNTAG was unique, and for 165 samples, both the UPTAG and DNTAG sequences were redundant with another sample where the called genotype was isogenic or highly similar (≤2 differences), strongly indicating the presence of a single strain in multiple wells. When processing the sequencing data, a single random strain was chosen to represent each unique UPTAG and DNTAG sequence.

**Examining Putative Wild-Type Pool Strains**

For 73 unique MAT**a** strains and 131 MATα strains, pooled genotyping had called the strains as wild-type. These strains were isolated from the pool and tested for growth in SC-Ura for confirmation. Of the 97 MATα strains, 74 exhibited no detectable growth in SC-Ura, and likely arose from remaining barcoder parents which had escaped a previous SC-Ura selection step. The remaining 23 strains were discarded. All 73 MAT**a** strains exhibited growth in SC-Ura. Individual genotyping was performed for 40 of these MAT**a** strains, and their stated genotype was corrected (Additional File XX). The other 33 strains were discarded from analysis. When calculating linkage and distribution of gene knockouts, the wild-type MATα strains were excluded from analysis as they are likely parental strains rather than progeny from the mating.

**Population Growth Profiling by High-Throughput Sequencing**

All random knockout progeny were combined into two separate liquid YPD pools, separated by mating type. The effects of 16 different drugs on strain growth were tested to find a concentration which inhibits overall growth by approximately 20%. The determined doses used in the pooled experiment are listed in Additional File XX. All drugs used were dissolved in XX% DMSO, which was used as a solvent control.

0.0625 OD600 nm from a YPD only pool (corresponding to approximately 1250 cells per strain) were added media containing drug/solvent. For the solvent control, a 0 generation sample was immediately harvested afterwards. After growth to 2 OD600 nm, a sample was taken from each drug and 0.0625 OD600 nm of cells were added to fresh medium. This process was repeated three times. Collected samples corresponded approximately to 5, 10, 15, and 20 generations of growth.

Harvested samples were subject to genomic DNA extraction. Two separate PCR reactions were performed on

**Sequence Data Processing**

**Finding Complex Genetic Interactions using a Linear Model**

**Targeted Mating and Selection to Obtain 32 Knockouts**

**Liquid Growth Assays**