Online Methods

Construction of segregating pools. To construct segregating pools, we employ the Synthetic Genetic Array (SGA) marker system [1,2]. In our cross, we use a BY parent that is MATα *can1*Δ*::STE2pr-SpHIS5 lyp1*Δ *his3*Δ*1* and an RM parent that is MAT**a** *AMN1*^{BY} *his3*Δ*0::NatMX ho::HphMX*. These strains were crossed and a diploid zygote was recovered.

To create the segregating pools, a single colony of the diploid progenitor was inoculated into 100 ml YPD and grown to stationary phase. The diploid culture was spun down and the supernatant was decanted. The diploid pellet was then resuspended in 200 ml Spo++ sporulation medium (http://www.genomics.princeton.edu/dunham/sporulationdissection.htm). The sporulation was kept at room temperature with shaking and monitored for the fraction of diploids that had sporulated. Once more than 50% of the diploids had sporulated, the culture was deemed suitable for downstream use.

The next step in the generation of segregating pools was to select for MAT**a** haploids. 50 ml of the sporulation were spun down and then the supernatant was decanted. The pellet was resuspended in 1 ml water. 300 ul β-glucoronidase (Sigma, Catalog #: G7770) were added to the preparation and the mixture was incubated at 30 °C for an hour. ~50 ul of glass beads (Sigma, Catalog #: G8772) were then added and the sample was vortexed for 2 minutes. The sample was incubated for an additional hour at 30 °C, followed by a second round of vortexing for 2 minutes. Water was added to the sample so that the total volume was 20 ml. The spore preparation was spread onto YNB +

canavanine/thialysine, with 100 ul of sample going onto each plate. The plates were incubated at 30 °C. MAT**a** haploids typically grew up after ~2 days.

The final step in pool creation was to mix together MAT**a** segregants selected on different plates. 10 ml water were poured onto a plate and a sterile spreader was used to remove the segregants from the plate. The cell mixtures from each plate were then pipetted off the plates into a separate container. The pool was spun down and the water decanted. For drug selections, the cell pellet was resuspended in 1.5 ml YPD per scraped plate. The segregant pool was incubated at 30 °C for an hour. 100 ul of this segregant pool was then spread onto each selection or control plate. For sorting of Mitotracker Red-stained cells, haploid segregants selected on YNB + canavanine/thialysine were scraped from plates and inoculated into YNB + canavanine liquid medium at a concentration of around ~3 x 10^6 cells/ml. The cells were grown for approximately three generations to a density of ~2 x 10^7 cells/ml.

Drug selections with segregating pools. X-QTL should be most powerful when selections are stringent, as this implies that one is enriching for segregants that are phenotypically extreme and are likely to possess multiple alleles that affect a trait in the same direction. For cell sorting, such selections are straightforward, as individual cells exhibiting a trait value within a specified range can be isolated. For chemical resistance mapping, achieving a stringent selection is slightly more challenging, as a whole population of cells is plated and one can only enrich for segregants with high trait values.

Drug selections with segregating pools require finding the optimal concentration to use for a particular compound prior to X-QTL mapping. To do this, we plate segregating pools across a range of concentrations. The concentration at which we start to resolve individual colonies on plates is the concentration that we use for X-QTL mapping. The fact that we observe ~10^2 to ~10^3 individual colonies when we plate more than 10^6 individuals implies that we are selecting far into the resistance tail of the phenotype distribution. Final concentrations used for the chemical selections are in Table S1. After selection was completed, several replicate selection plates were scraped, pooled, and frozen at -80 °C. Control experiments were also conducted by plating segregating pools on YPD without any drug added, and pools were collected and stored in the same manner as the selections.

We attempted to combine MAT**a** selections with our chemical resistance selections by incorporating a chemical of interest into our YNB + canavanaine/thialysine plates. We found that this approach worked far worse than separating the selection of MAT**a** haploids and the selection of resistant segregants into two steps.

Microarray description. We designed our array using 21,994 BY and RM allele-specific probe pairs. These pairs cover 17,566 SNPs that differentiate BY and RM, at an average spacing of one marker every ~700 bp. The BY-specific probes were designed as part of a separate study of optimal probe design parameters for DNA genotyping arrays and were chosen to minimize the variance in T_m values across probes [3]. For this study, we used the previously designed BY-specific probes and made an additional probe specific to the RM sequence. In order to maximize the sensitivity of our genotyping

array, probes were chosen to have the interrogated SNP within the middle five bases of a given probe. Our custom two-color microarray was manufactured by Agilent.

Comparison of microarray data to sequencing data. DNA from the same control and 4-NQO-selected segregating pools was hybridized to the microarray and sequenced on the Illumina Genome Analyzer using 75 bp reads. Two biological replicate control and two biological replicate 4-NQO-selected pools were sequenced. Except for one of the replicate controls that was sequenced in a single lane, each sample was sequenced in two lanes. To analyze the Genome Analyzer data, sequencing reads were mapped to the BY genome using ELAND and the Illumina EXPORT files were converted into SAM format using SAMTOOLS [4]. The PILEUP function in SAMTOOLS was used to reformat the sequence data. Sequence data at polymorphic sites included on the genotyping microarray were extracted from the PILEUP file and only these sites were analyzed. The polymorphic sites were subjected to a quality filter, with only sites having a quality score of 10 or higher used. The coverage was ~60X per site in each lane. Figure S3 shows only one lane (~60X) of sequence data from a 4-NQO selection, four aggregated lanes of sequence data (~240X) from both 4-NQO selections, and a single microarray. Even at 60X sequencing coverage, peaks are discernible, although the variance in measured allele frequencies is high. 240X coverage provides results comparable to the genotyping microarray. Our results suggest that both X-QTL and X-QTL-seq are useful approaches to genetic mapping in pools of cross progeny.

Mapping results for drug traits. Prior to analyses, each array was subjected to a

quality check that both allele-specific probes for a given probe set had successfully hybridized. Bad probe sets were excluded from downstream analyses. We conducted separate analyses for the drug selection and FACS-based selection experiments.

For the drug selections, t-tests were conducted comparing results from two independent selection experiments to results from 13 independent control experiments. T-tests were conducted with the variances of the two groups set to be equal. The $-\log_{10}(p)$ values were then used for unsupervised peak calling. We found that an approach that scanned the genome for inflection points in the slope of the average $-\log_{10}(p)$ values worked best. By definition, a peak is a point at which the slope of the data changes sign. When scanning $-\log_{10}(p)$ values, which are always positive, a peak is represented by a positive to negative sign change.

To identify inflection points, we first smoothed the data by averaging the $-\log_{10}(p)$ values within 50 kilobase sliding windows. We then scanned the genome chromosome-by-chromosome by resistance trait using sliding window linear regression. We fit linear regressions over 100 kilobase sliding windows and used the slope of these regressions to estimate the locations of peaks. A special case was allowed at the ends of chromosomes in which peaks were recorded if the slope was negative at the top of the chromosome or positive at the bottom of the chromosome. The average $-\log_{10}(p)$ value at estimated peaks was recorded and used for thresholding. The same approach was used to analyze 1,000 permutations of the chemical resistance dataset, in which two randomly chosen arrays ("selections") were compared to 13 randomly chosen arrays ("controls"). A requirement was set on the permuted datasets that the selection arrays

never be biological replicates of the same real trait selection. Because of uncertainty about what constitutes a distinct peak under cases of close linkage, we set a requirement that two peaks could not occur within 200 kilobases of each other. Increasing or decreasing this proximity threshold results in a slightly different number of called peaks, but does not affect the general findings of the paper. Inflection points detected in the permutations were used to set an empirical FDR threshold of 0.05. We used a global FDR threshold, as opposed to a trait-level FDR, as most observed expected-observed peak relationships at the trait-level were very close to the global relationship (Figure 3D). Average $-\log_{10}(p)$ plots, as well as significant peaks, are provided for each trait in Figures S5A-Q.

For the FACS experiment, three low, three high, and three whole population biological replicates were generated. Because of the small number of arrays in the experiment, permutations were unlikely to be useful for setting an empirical FDR threshold. Furthermore, because the data structure of the FACS experiment, which used two tails of the segregant distribution, was different from the drug selections, which used only one tail of the segregant distribution, we could not utilize the drug selections in permutations of the FACS data. For these reasons, we used QVALUE [5], which estimates the FDR using the distribution of p-values in an experiment, to determine probes that were significant at an FDR of 0.05. We show this threshold in Figure 4.

All analyses were conducted in R.

Online Methods References

[1] Tong, A.H. *et al.*, Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294 (5550), 2364-2368 (2001).

[2] Tong, A.H. & Boone, C., High-throughput strain construction and systematic synthetic lethal screening in *Saccharomyces cerevisiae* in *Yeast Gene Analysis - Second Edition. Methods in Microbiology*. (Elsevier Ltd., 2007), Vol. 36, pp. 369-707.

[3] Gresham, D. *et al.*, Optimized detection of sequence variation in heterozygous genomes using DNA microarrays with isothermal-melting probes. *PNAS* (in press).

[4] Li, H. *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25 (16), 2078-2079 (2009).

[5] Storey, J.D. & Tibshirani, R., Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100 (16), 9440-9445 (2003).