Parallel Genetic Changes and Nonparallel Gene–Environment Interactions Characterize the Evolution of Drug Resistance in Yeast

Aleeza C. Gerstein,[1] Dara S. Lo, and Sarah P. Otto
Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

ABSTRACT Beneficial mutations are required for adaptation to novel environments, yet the range of mutational pathways that are available to a population has been poorly characterized, particularly in eukaryotes. We assessed the genetic changes of the first mutations acquired during adaptation to a novel environment (exposure to the fungicide, nystatin) in 35 haploid lines of *Saccharomyces cerevisiae*. Through whole-genome resequencing we found that the genomic scope for adaptation was narrow; all adapted lines acquired a mutation in one of four late-acting genes in the ergosterol biosynthesis pathway, with very few other mutations found. Lines that acquired different ergosterol mutations in the same gene exhibited very similar tolerance to nystatin. All lines were found to have a cost relative to wild type in an unstressful environment; the level of this cost was also strongly correlated with the ergosterol gene bearing the mutation. Interestingly, we uncovered both positive and negative effects on tolerance to other harsh environments for mutations in the different ergosterol genes, indicating that these beneficial mutations have effects that differ in sign among environmental challenges. These results demonstrate that although the genomic target was narrow, different adaptive mutations can lead populations down different evolutionary pathways, with respect to their ability to tolerate (or succumb to) other environmental challenges.

POPULATIONS adapt to stressful environments through the fixation of beneficial alleles. The number of advantageous mutations accessible to a population within one or a few mutational steps ["mutational neighborhood" (Burch and Chao 2000)] remains particularly poorly characterized, especially in eukaryotes. This is an important factor, however, as the number of mutations in concert with their pleiotropic effects will directly influence the range of evolutionary pathways available to different populations. The first beneficial mutations to fix are of particular interest, as genetic and gene–environment ($G \times E$) interactions may dictate the fixation of subsequent mutations. Knowledge of the number of available pathways may help us predict whether two populations subjected to similar selective pressures in allopatry might accumulate and fix different mutations. If this frequently occurs, reproductive isolation could evolve purely by chance fixation of different mutations [the mutation-order hypothesis (Schluter 2009)]. We thus sought to determine the mutational neighborhood of adaptive mutations in one environment and to characterize the pleiotropic effects of these mutations to different environmental challenges.

A fruitful approach to characterize the genotypic basis of adaptation has involved experimental microbial studies, where multiple replicate populations are initiated with the same ancestral culture and evolved under the same conditions for several generations (Conrad *et al.* 2011). Targeted resequencing of specific genes in replicate populations evolved for hundreds or thousands of generations at large population size (where selection should overwhelm drift) has demonstrated that in many cases the same genes repeatedly acquire mutations (Cooper *et al.* 2003; Pelosi *et al.* 2006; Woods *et al.* 2006; Ostrowski *et al.* 2008; Barrick *et al.* 2009). An examination of diverse clinical isolates of *Pseudomonas aeruginosa* has also repeatedly implicated the same genes during the acquisition of resistance to quinolone

doi: 10.1534/genetics.112.142620
Manuscript received February 8, 2012; accepted for publication June 11, 2012
Supporting information is available online at http://www.genetics.org/content/suppl/2012/06/19/genetics.112.142620.DC1.
[1]Corresponding author: Department of Zoology and Biodiversity Research Centre, 6270 University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.
E-mail: gerstein@zoology.ubc.ca

(Wong and Kassen 2011). Parallel genotypic evolution is not restricted to the utilization of single genes, as parallel transposition mutations (Chou *et al.* 2009) and large-scale aneuploid events (Selmecki *et al.* 2009) have also been documented in replicate lineages evolved under the same conditions. Furthermore, the magnitude of genetic parallelism has been shown to be influenced by the selective environment (Anderson *et al.* 2003; Gresham *et al.* 2008), depending on both the size of the genomic target for beneficial mutations and the probability of establishment of different mutations.

Targeted resequencing studies may, however, paint a skewed picture of the extent to which parallel mutations underlie evolutionary change, as only a few genes are typically examined and such candidate genes may be more likely to be repeated targets of beneficial mutations. Only through whole-genome resequencing (WGS) can the full array of beneficial mutations and their chance of appearing repeatedly be assessed. Over the last few years, a broader picture of the types of mutations acquired in long-term experimental evolution lines has been painted by WGS. The results suggest that while the number of different genes available to adaptation and the types of mutations depend on both the species and the environment, some generalities from targeted resequencing studies seem to hold. In all cases a small number of genes have been the target of beneficial mutations in independently evolved lines, with nonsynonymous single-nucleotide changes (SNPs) being the most common type of mutation (Herring *et al.* 2006; Araya *et al.* 2010; Kishimoto *et al.* 2010; Minty *et al.* 2011; Toprak *et al.* 2011; Tenaillon *et al.* 2012). Considerable variation among experiments is found in the total number of different genes targeted, however, and variation is also present for the absolute frequencies of different classes of mutations (*e.g.*, copy number variants, insertions, deletions, and regulatory changes).

The order of mutational steps can have a tremendous impact on the fitness effect of subsequent mutations (Weinreich *et al.* 2005), both in magnitude and in sign, implying that the first adaptive step taken can alter the path of evolution. While WGS has allowed us to leap forward in our understanding of the genetic basis of adaptation, fewer WGS studies have focused on the first step of adaptation, and none have yet characterized the first steps in a eukaryote. The data that exist in viruses and prokaryotes suggest that the first mutations to be selected also tend to be clustered in relatively few genes. Rokyta *et al.* (2005) identified 10 unique nonsynonymous single-step mutations through WGS within two different viral genes of ΦX174. Similar results were found through targeted resequencing in both *P. aeruginosa* [where 15 unique mutations were identified in *rpoB* (MacLean and Buckling 2009)] and *P. fluorescens* [five nonsynonymous SNPs were found in *gyrA* and four mutations were found within three efflux pump regulatory sites (Bataillon *et al.* 2011)]. As with all targeted resequencing studies, it remains unknown how many additional mutations were present in the lines acquired in *Pseudomonas*. It has also not yet been determined whether eukaryotes accumulate mutations in a similar (*i.e.*, largely parallel) manner. A number of fundamental characteristics differ between prokaryotic and eukaryotic genomes [*e.g.*, chromosome structure, the number of replication origins, the amount of noncoding DNA, and the degree of transcript processing, just to name a few (Poole *et al.* 2003)]. These or other factors could affect the nature of mutations acquired under stressful conditions in eukaryotic genomes.

To assess the mutational neighborhood allowing adaptation to a novel stressful environment in a eukaryote we developed an assay to isolate multiple adapted lines of haploid *Saccharomyces cerevisiae*. We exposed 240 replicate lines initiated with ~100,000 progenitor cells to a level of stressor that inhibits growth of the ancestral strain. By immediately isolating cells that were able to grow in this environment, we limited the number of mutational hits in the genome and reduced the potential influence of clonal interference. Through WGS we pinpointed the genetic basis of adaptation for each lineage. We chose a polyene antibiotic, nystatin, as the stressful environment in which to acquire mutations. Nystatin binds to ergosterol (the primary sterol in the fungal membrane) to form porin channels that increase membrane permeability and allow cellular components (including potassium ions, sugar, and metabolites) to leak out of the cell (Carrillo-Munoz *et al.* 2006; Kanafani and Perfect 2008). The resulting change in potassium concentration leads to an osmotic imbalance between the vacuole and cytoplasm and an enlarged vacuole (Bhiyan *et al.* 1999). Transcriptional profiling has identified membrane transporters and the cell stress response as the major cellular components affected by exposure to nystatin (Hapala *et al.* 2005). Previous work has identified resistant mutants in *Candida albicans* and *S. cerevisiae* that show defects in genes involved in the ergosterol biosynthesis pathway, with mutants exhibiting an altered sterol content in the cell membrane (Bhiyan *et al.* 1999; Ghannoum and Rice 1999; Kanafani and Perfect 2008). Whether the ergosterol pathway would always be involved in resistance evolution, or whether beneficial mutations could be recruited from other membrane components or from altered ion pumps, remains unknown. The first goal of our research was thus to document the genetic basis of adaptation to nystatin in many replicate lineages and to measure the fitness benefit gained by each adapted lineage.

We then determined whether the mutations we identified exhibited different responses to other stressful environments ($G \times E$). Given that populations often face multiple environmental challenges simultaneously, the scope for adaptation would be greatly reduced if adaptive mutations in the presence of a single environmental change always exhibit reduced tolerance to other environmental challenges (strict trade-offs). The idea that strict trade-offs should exist during adaptation to novel environments is long-standing and multiple hypotheses have been put forward to explain the

physiological basis of trade-offs (Pörtner *et al.* 2006), yet trade-offs are not universally found (Ostrowski *et al.* 2005; Bennett and Lenski 2007; Hereford 2009). Another possibility is that mutations vary in the subset of environments in which they are beneficial [*i.e.*, there is "sign $G \times E$," by analogy with sign epistasis (Weinreich *et al.* 2005)]. In this case, some mutations may be simultaneously beneficial to multiple types of change in the environment, allowing the organism to adapt more readily to complex environmental challenges. Furthermore, such gene–environment interactions imply that lineages carrying different first-step mutations would find themselves at different locations on the adaptive fitness surface after further changes in the environment. To explore the nature of gene–environment interactions among single adaptive mutations, we conducted a set of experiments to measure the fitness effects under different stressful conditions of mutations whose genetic basis is known, allowing us to compare mutations in different genes and at different sites within a gene. To explore a variety of environmental challenges, we varied levels of copper, ethanol, and salt, measuring growth of each line in each environment.

When exposed to nystatin, we found strong parallelism in the adaptive mutations that appeared within our *S. cerevisiae* lines at both the gene and the pathway levels, with only a limited number of genes being involved in the first step of adaptation. Mutations in different genes had significantly different fitness effects across environments, with some lines showing increased tolerance and others decreased tolerance to other stressors (sign $G \times E$). Our results thus provide support for the mutational-order hypothesis that adaptation to one environmental challenge may well drive isolated populations down different evolutionary pathways, with significant differences in starting fitness when faced with further environmental challenges.

Materials and Methods

Strain background and mutation acquisition

Mutations were acquired in haploids of genotype BY4741 (*MATa his3△1 leu2△0 met15△0 ura3△0*) derived from S288C. Stocks of the ancestral culture and relevant gene deletion lines (erg6Δ, erg3Δ, and erg5Δ, see below) were ordered from Open Biosystems (Thermo Fisher Scientific) and streaked on a YPD plate, and a single colony was isolated and frozen. This single colony from BY4741 served as the progenitor culture for all experiments, and we refer to this as the ancestral strain. Mutation accumulation was carried out in 1-ml culture in 96-deep–well plates (2-ml polypropylene plates with a conical bottom), with shaking at 200 rpm at 30°. We acquired mutations in two screens, separated by 2 weeks. To initiate each screen, the ancestral strain was streaked onto a YPD plate from the frozen isolate and grown for 48 hr. A single random colony was then picked and grown for 24 hr in 10 ml YPD at 30°, with shaking at 200 rpm.

We initiated mutation acquisition in the first screen by transferring 10 μl of the ancestral overnight culture into 1 ml of YPD + 4 μM nystatin into the 60 inner wells of a 96-deep–well plate. The second screen was identical except we initiated 180 replicates into the inner wells of three 96-deep–well plates. The level of nystatin was determined in preliminary experiments as the level that showed only sporadic growth of the ancestral strain within 2–7 days of incubation, suggesting that growth required a mutational event (results not shown). Growth in nystatin for each screen was checked and recorded daily by visual examination of the bottom of the 96-well plates. A small amount of growth would typically be observed one day, with full growth on the second or third day (where full growth is approximately equivalent to the amount of turbidity and precipitate seen in the ancestral strain after 24 hr of growth in YPD). Occasionally full growth took up to 4 days. Each well that showed growth (even slow growth) was marked as a "putative mutation" line.

On the first day that full growth was recorded for each putative mutation line, the well was thoroughly mixed by pipetting and culture was streaked onto a YPD plate. Putative mutation lines were obtained in this way from 64 of the 240 inoculated wells, with no growth observed in the remaining wells over 7 days. After 48 hr of growth on the plate, we visually assessed each line for petite mutations (mutations that affect mitochondrial function and prevent respiration; these colonies present as much smaller than normal on a YPD plate). About half of the putative mutation lines showed evidence of petite mutations. All suspected petites were confirmed by lack of growth on a YPG plate, a medium that requires respiration for growth, and these lines were discarded to focus on nuclear mutations. For each of the remaining 35 lines, eight colonies were haphazardly picked off the YPD plate, placed back into 8 wells containing 1 ml of YPD + 4 μM nystatin, and assayed for growth. Often all eight colonies picked would exhibit similar growth patterns, but sometimes colonies varied with respect to the number of days to full growth or even whether any growth was observed in nystatin within 48 hr. To avoid analyzing nonmutant cells that might have been segregating, we randomly picked a single well that showed any level of growth in nystatin for each mutation line. The 1-ml culture from this well was mixed with 1 ml 30% glycerol and frozen; this freezer culture constitutes the material for all future experiments. These 35 lines were labeled BMN1–35 (BMN: "beneficial mutation nystatin"), as described in the Table 1 legend.

Sequencing

Freezer culture from each BMN line was streaked onto YPD plates and grown for 48 hr. A single colony was then haphazardly picked for each line and grown for 24 hr in YPD. DNA was extracted (Sambrook and Russell 2001) and sequenced in 100-bp single-end fragments, using Illumina's (San Diego) HiSeq 2000. Library preps followed standard

Table 1 The genetic basis of BMN line ergosterol mutations

Line (BMN)	Gene	Genome position (Chr.bp)	Position in gene (in nucleotides)	Mutation	Amino acid change
1	*ERG7*	VIII.241,194	2097	C > G[a]	Phe699Leu
2–4	*ERG6*	XIII.252,861	130	C > T[a]	Gln44Stop
5	*ERG6*	XIII.252,772	119	C > A[a]	Tyr73Stop
6	*ERG6*	XIII.252,723	268	C > A	His90Asn
7–10	*ERG6*	XIII.252,612	379	G > C[a]	Gly127Arg
11–15	*ERG6*	XIII.252,596	395	C > –[a]	
16	*ERG6*	XIII.252,349	642	G > C	Leu214Phe
17–20	*ERG6*	XIII.252,322	669	C > G	Tyr223Stop
21	*ERG3*	XII.254,047	187	A > T[a]	Arg63Stop
22	*ERG3*	XII.254,087	227	C > A	Ser76Stop
23	*ERG3*	XII.254,144	284	C > A[a]	Ser95Stop
24–27	*ERG3*	XII.254,475	615	G > A	Trp205Stop
28	*ERG3*	XII.254,501	641	29-bp duplication[b]	
29	*ERG3*	XII.254,516	656	G > A	Trp219Stop
30	*ERG3*	XII.254,563	703	G > A	Gly235Ser
31	*ERG3*	XII.254,757	897	C > A	Tyr299Stop
32	*ERG3*	XII.254,758	898	G > C[a]	Gly300Arg
33	*ERG3*	XII.254,780	920	A > C	Asp307Ala
34	*ERG3*	XII.254,840	980	A > −	
35	*ERG5*	XIII.301,120	253	60-bp deletion[c]	

Each BMN line carried a single mutation in one of four genes at the end of the ergosterol pathway. We numbered each line sequentially based on the location of each ergosterol mutation. Lines with mutations in genes farthest in the pathway from producing ergosterol (the end product) have lower numbers; within a gene, mutations nearer the start codon were given lower numbers.

[a] Mutations were confirmed with Sanger sequencing (when multiple lines shared a mutation, we confirmed the mutation in only a single line: BMN3, BMN9, and BMN13).

[b] Confirmed with Sanger sequencing: base pairs 641–669 are duplicated and inserted after base pair 669.

[c] Confirmed with Sanger sequencing: deletion of 60 bases between base pairs 253 and 312.

Illumina protocols (2011 Illumina, Inc., all rights reserved), with 12 uniquely barcoded strains run together on a lane. The resulting genomic sequence data were processed using Illumina's CASAVA-1.8.0. Specifically, configureBclToFastq.pl was used to convert to fastq and separate the sequences by barcode (allowing one mismatched base pair). ConfigureAlignment.pl (based on the alignment program ELAND) was then used to align each sequence to the yeast reference genome (scergenome.fasta downloaded from the Saccharomyces Genome Database, http://downloads.yeastgenome.org/genome_release/r64/).

SNPs and small insertions and deletions (indels) were then called using configureBuild.pl. Average coverage per mapped site across the strains (excluding mitochondrial genes) was 44.0 (with a minimum average coverage per site of 4.7 for BMN13). Data on relative coverage per chromosome are illustrated in Supporting Information, Figure S1, which indicate that one line (BMN27) had an additional copy of chromosome 2 (denoted as a black circle in Figure S1). Custom UNIX and perl scripts were then used to parse the output files. Illumina data from an independent project using the same ancestral strain were used to identify mutations that were common to the ancestor, and all such mutations were ignored. Given that our initial lines were haploid, variants identified by configureBuild.pl as heterozygous were also discarded as likely alignment or sequencing errors. Similarly, variants involving repeat elements were discarded. All remaining variants were checked in the alignments, using tview in samtools-0.1.7a (Li *et al.* 2009); variants that were not supported by multiple fragments starting from different positions were also discarded (typically near deletions or gaps in the alignment). Finally, the same procedure was repeated, but using the bwa software package to perform the alignment (Li *et al.* 2009) along with samtools-0.1.7a to identify SNPs (Li *et al.* 2009), using the -bq 1 option to limit data to reliable alignments. All SNPs (Table 1 and Table S1) were identified using both methods.

Sequence alignments were also manually checked using tview in samtools-0.1.7a for each of the four genes harboring beneficial mutations (*ERG3*, *ERG5*, *ERG6*, and *ERG7*) to look for larger rearrangements or other changes not identified by the above procedure. Two additional large-scale mutations were identified from gaps in the alignments. To determine the nature of these rearrangements, the fastq files containing the unaligned short-read sequences were directly searched for sequences on either side of the alignment gap, confirming a 60-bp deletion in *ERG5* within line BMN35 and a 29-bp duplication in *ERG3* within line BMN28 (Table 1). For confirmation, we Sanger sequenced the appropriate gene from BMN lines representing 10 of the 20 unique mutations found (Table 1). In all cases the Sanger sequence data matched our analysis of the genomic sequence data. It is worth noting repetitive sequences were often not mapped and were ignored when mapped, so that any mutations in such sites or any larger-scale rearrangements not involving these four genes would likely have been missed. Nevertheless, 95.3% (SE = 0.15%) of the known sites in the *S. cerevisiae* reference genome had coverage in our

alignments, so that we are likely to have detected most nucleotide mutations and small indels that occurred.

Sterol assay

We compared the sterol profile of the ancestral strain (BY4741) and BMN lines, using a spectrophotometry-based assay. When more than one BMN line shared the same ergosterol mutation, we randomly chose one line to represent that group. Sterols were extracted using the alcoholic potassium hydroxide method as previously described (Arthington-Skaggs *et al.* 1999). BMN culture streaked to a single colony on a YPD plate was inoculated into 50 ml of YPD and grown at 30° for 48 hr at which point the optical density (at 630 nm) was measured to record cell concentration. Cells were then harvested by centrifugation at 2700 rpm for 5 min and washed twice with sterile distilled water. Three milliliters of 25% alcoholic KOH was added to each pellet and vortexed for 1 min. The sample was then incubated in an 80° water bath for 1 hr and then cooled to room temperature. To extract the sterols, 1 ml of sterile distilled water and 3 ml of heptane were added and vortexed for 3 min. A 200 μl aliquot of the heptane layer was added to 800 μl of 95% ethanol, and the absorbance was immediately read every 3 nm between 200 and 300 nm with a Thermo BioMate 3 spectrophotometer.

Nystatin tolerance

To determine the breadth of nystatin tolerance conferred by each mutation, a growth assay was performed to measure the half-maximal inhibitory concentration (IC_{50}) of nystatin. Freezer stock from each BMN line was inoculated into 1 ml of YPD in one well of a 96-deep–well plate and grown for 48 hr. To standardize the starting density of cells, the optical density (OD) of 200 μl from each well was measured using the BioTek microplate reader (BioTek Instruments, Winooski, VT) and diluted to the sample with the lowest OD (usually between 0.7 and 0.9). Two hundred microliters of the standardized culture was then added to 400 μl of YPD to obtain the final volume necessary for the assay inoculations. For each BMN line, 12 μl of the dilute culture was then inoculated into a well containing 1 ml YPD plus 1 of 10 levels of nystatin (0 μM, 2 μM, 4 μM, 8 μM, 12 μM, 16 μM, 46 μM, 96 μM, 116 μM, and 200 μM), each replicated in 4 different nonadjacent wells. Plates were sampled at 72 hr to measure OD. Wells were manually mixed and 150 μl aliquots were taken from each well and read on the BioTek reader.

A maximum-likelihood model was fit to the data to determine IC_{50}. The logistic function

$$y = \frac{y_{\max}\exp(a(x - \mathrm{IC}_{50}))}{1 + \exp(a(x - \mathrm{IC}_{50}))} + N(0, \sigma) \qquad (1)$$

was used, where x represents the tested concentration of nystatin, y represents the observed OD following 72 hr of growth, and $N(0, \sigma)$ represents a normal deviate with mean zero and standard deviation, σ. The fitted parameters were $y_{\max}$ (the maximal OD under full growth), IC_{50} (the nystatin concentration at which OD is half maximal), a (the slope of the logistic curve at $x = IC_{50}$ divided by $y_{\max}/4$), and σ. Prior to fitting the data using this likelihood procedure, all nystatin concentrations were ln transformed (so that percentage changes, not absolute differences, in nystatin matter), although we report all values of nystatin concentration and IC_{50} on the original scale. The maximum-likelihood point was found in R, using the subplex method of optim, as implemented in the find.mle routine of the diversitree package (FitzJohn *et al.* 2009). The find.mle routine allows lower and upper limits to the parameters to be specified in the search routine (we used lower, $y_{\max} = 0.8$, $IC_{50} = 0.0000001$, $a = -50$, and $\sigma = 0$; and upper, $y_{\max} = 1.2$, $IC_{50} = 116$, $a = 0$, and $\sigma = 10$; lower and upper $y_{\max}$ and IC_{50} were based on observations).

To determine whether the IC_{50} of a mutant line was significantly different from that of the ancestral strain, a likelihood model was fitted to the data from the mutant line and the ancestral line, allowing each of these two lines to have its own values of $y_{\max}$, IC_{50}, a, and σ. This "full" model was then compared to a constrained model where $IC_{50,\mathrm{mutant}} = IC_{50,\mathrm{ancestral}}$, using a likelihood-ratio test. If the drop in log-likelihood between the full and the constrained model was $> \chi^2_{1,0.05}/2 = 1.92$, we rejected the hypothesis that IC_{50} was the same for the two lines.

Fitness proxies in a permissive environment and the evolutionary environment

We assessed growth in both the evolutionary environment (YPD + 4 μM nystatin) and an unstressful environment (standard laboratory YPD), using two fitness proxies. We measured the maximal growth rate to measure how quickly yeast cells are taking in nutrients and growing during the exponential phase of growth and optical density at 48 hr to capture the ability to turn resources into biomass. Both fitness proxies were determined using previously described methods (Gerstein and Otto 2011) that utilize the Bioscreen C Microbiological Workstation (Thermo Labsystems), which measures OD in 100-well honeycomb plates. In brief, plates were streaked from frozen stock onto YPD plates for all lines and the ancestral strain and allowed to grow for 72 hr. Inoculations containing one colony (haphazardly chosen) for each BMN line and five separate inoculations of the ancestral strain (each from a different single colony) were then allowed to grow for 48 hr in 10 ml YPD. One hundred microliters was transferred into 10 ml of fresh YPD and mixed thoroughly, and four 150 μl aliquots for each line were placed into nonadjacent bioscreen wells. The bioscreen plates were grown at 30° for 48 hr with constant shaking; OD readings were automatically taken every 30 min. We determine the maximal growth rate for each well as the spline with the highest slope, from a loess fit through ln-transformed optical density data, using an analysis program written by Richard FitzJohn in R (R Development Core

Team 2011). OD at 48 hr was used as a second fitness measure (OD48). As can be seen from the raw growth curves (Figure S2), the lines have stopped growing by 48 hr in YPD, and this measure thus reflects efficiency (*i.e.*, ability to turn nutrients into cellular material). In nystatin, by contrast, some lines may still be growing, and this assay thus represents a combined measure of growth rate and efficiency.

Assessing gene–environment interactions

The ecological tolerance (measured as IC_{50}) was determined for each line in copper ($CuS0_4$), ethanol, and salt (NaCl). The tolerance assays in these environments were conducted as previously described for nystatin; we measured growth after 72 hr in eight levels of copper (0 mM, 1 mM, 2 mM, 4 mM, 6 mM, 8 mM, 10 mM, and 12 mM), seven levels of ethanol (0%, 2%, 4%, 6%, 8%, 12%, and 14%), and eight levels of salt (0.2 M, 0.4 M, 0.6 M, 0.8 M, 1.2 M, 1.4 M, 1.6 M, and 2 M). These levels were chosen based on preliminary data that indicated the approximate position of IC_{50}. Significance was determined as with tolerance to nystatin, using a maximum-likelihood test that compared a model fit with two IC_{50} parameters (one for the mutation line and one for the ancestor) to a model with only one IC_{50} value.

Results

We acquired 64 haploid lines of *S. cerevisiae* that were resistant to nystatin in two independent screens. We eliminated all lines that could not respire, which left us with 35 lines that we term BMN lines. Through WGS with the Illumina platform, we found that each line carried a single mutation in one of four late-acting genes in the ergosterol biosynthesis pathway (Figure 1). Within these lines, we found 1 line each with mutations in *ERG7* ("BMN-*erg7*") and *ERG5* ("BMN-*erg5*"), 7 unique mutations in *ERG6* within 19 lines (collectively referred to as "BMN-*erg6* lines"), and 11 unique mutations within 14 lines in *ERG3* ("BMN-*erg3* lines"). We recovered multiple classes of mutations including nonsynonymous SNPs, premature stop codons, small indels (<3 bp), one 60-bp deletion, and one 29-bp duplication (Figure 1); the precise nucleotide and amino acid changes as well the numbering scheme for BMN lines are presented in Table 1.

For five mutations, the same sequence change was observed in multiple lines (Table 1). There are three potential explanations for this finding. The most likely is that mutations were initially segregating in the source population before the stressor was applied (see File S1 and File S2). A second explanation, that contamination occurred among wells in the 96-well plates, is possible, but fails to explain the similar timing of appearance of identical hits (Table S1). The independent appearance of the same sequence changes is also possible, but in no case did we observe the same mutation in the two screens (Table S1). While we treat each line as independent for statistical purposes, we note that combining lines with the same ergosterol mutation leads to the same conclusions (see File S1).

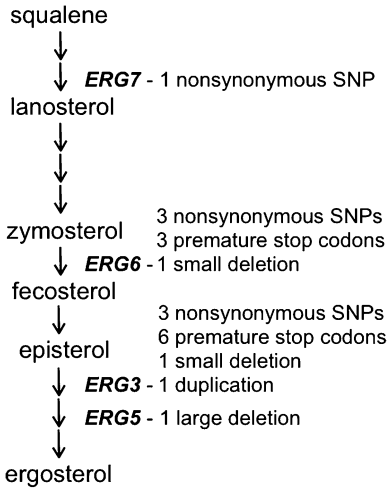


Figure 1 Twenty unique mutations were found in four late-acting genes in the ergosterol biosysnthesis pathway. Each arrow represents one gene in the pathway that converts squalene to ergosterol. See Table 1 for detailed information on the genetic nature of each mutation.

We did not expect to see many mutations other than those conferring a fitness benefit given the relatively small initial population size ($\sim 10^5$ cells), small genome-wide mutation rate (Lynch *et al.* 2008), and short time frame of the experiment (we stopped once growth could be observed, thus minimizing the number of generations; File S2). Nevertheless, we identified a small number of additional point mutations (five synonymous, eight nonsynonymous, and one nonsense changes in genes that are not part of the ergosterol pathway and five mutations in intergenic regions; Table S2). The majority of these mutations were unique to a single BMN line, but two mutations were found in multiple lines. A nonsynonymous change from glutamic acid to lysine in *FCY2* was found in four lines (BMN24–27), and a synonymous mutation was found in *GDA1* in five lines (BMN11–15); in both cases, these two sets of lines also shared a primary ergosterol mutation (Table 1), strongly suggesting that these two sets may be derived from the same mutations that arose in the precursor population. The genome sizes of all lines were measured using flow cytometry, and no deviations from haploidy were found. Examining the depth of coverage from Illumina data (see *Materials and Methods*) uncovered one case of chromosomal aneuploidy (BMN27 had a duplicated chromosome II, Figure S1). We did not find strong evidence that any of the nonergosterol mutations influence fitness in the environments measured (see File S1), and thus we focus our discussion on the ergosterol mutations.

We first measured the sterol profile of all lines. This assay takes advantage of the characteristic four-peak curve produced by ergosterol and the late sterol intermediate 24(28) dehydroergosterol (DHE) that are present in wild-type cells

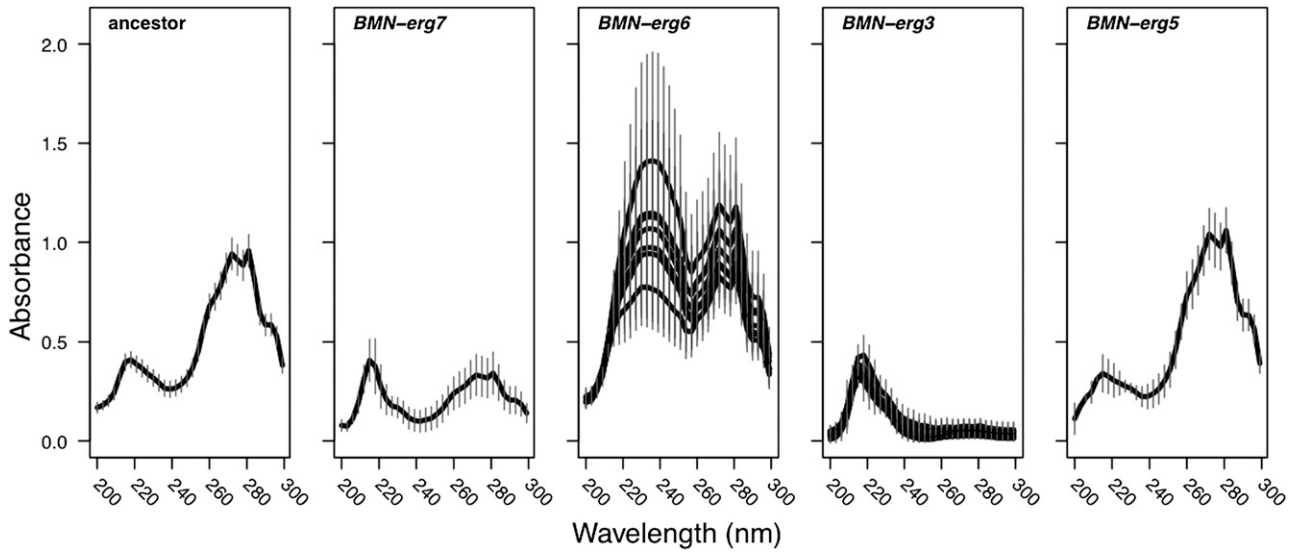


Figure 2 The sterol profile of all BMN lines is different from the ancestral profile except for the line with a mutation in *ERG5*. Sterol profile of each line was measured using a spectophotometry-based assay. For BMN lines that carry the identical ergosterol mutation, a single line was randomly chosen to represent the group (BMN-*erg6*: BMN3, –6, -9, -12, and -16–18; BMN-*erg3*: BMN25, -27, -29, and -30–33). Error bars depict the standard error of replicates measured on 3 different days.

(Arthington-Skaggs *et al.* 1999). All lines that carried mutations in the same ergosterol gene showed nearly identical sterol profiles (Figure 2). Interestingly, only BMN-*erg5* (the line with a mutation in the gene closest to the end of the pathway) had a sterol profile similar to that of the ancestral strain. The sterol profiles for BMN-*erg6* lines and BMN-*erg3* lines have a similar shape to previously published results of *erg3*Δ and *erg6*Δ obtained using the same protocol (Jensen-Pergakes *et al.* 1998; Mukhopadhyay *et al.* 2002). Surprisingly, however, our own measures of *erg6*Δ and *erg3*Δ (and *erg5*Δ) recovered the ancestral sterol phenotype (not shown).

All mutation lines had a significantly higher tolerance to nystatin than the ancestral strain, and many lines could tolerate nystatin at much higher levels than the 4 μM concentration used to isolate beneficial mutations (Figure 3). We measured the breadth of tolerance as IC_{50}, *i.e.*, the inhibitory concentration of the drug that reduced growth by 50%. The significance of changes in IC_{50} relative to that in the ancestor was determined by likelihood-ratio tests (Table S3). Replicate lines that carried different mutations in the same ergosterol gene showed similar IC_{50} values (Figure 3). A two-way ANOVA found that IC_{50} in nystatin has a very strong association with the ergosterol gene bearing a mutation ($F_3 = 252.4$, $P < 0.0001$) but was not affected by either the class of mutation within a gene (*i.e.*, nonsynonymous SNP, premature stop codon, or indel) or their interaction (mutation type, $F_2 = 0.92$, $P = 0.41$; interaction, $F_2 = 0.66$, $P = 0.53$). We then compared the tolerance of our lines to that of *S. cerevisiae* strains that carry gene knockouts for *ERG6*, *ERG3*, and *ERG5* (*erg7*Δ is inviable and could not be tested). Although all gene knockout lines did show increased nystatin tolerance compared to the ancestor (Figure 3), we found that while BMN-*erg5* and *erg5*Δ had similar IC_{50} values (Figure 3), BMN-*erg6* lines had a significantly higher nystatin tolerance than *erg6*Δ, and all but two BMN-*erg3* lines had a significantly lower nystatin tolerance than *erg3*Δ.

We also measured two fitness-related proxies for all lines in both the evolutionary environment (YPD + 4 μM nystatin) and an unstressful environment (standard laboratory YPD). When grown in nystatin, all BMN lines reached a higher optical density at 48 hr (OD48, Figure 4A) and had a higher maximal growth rate (Figure 4C) than the ancestral strain (significance determined by a *t*-test compared to five ancestral colonies, Figure 4 and Table S4, Table S5, Table S6, and Table S7). Conversely, the ancestor performed better than all BMN lines in YPD for both fitness proxies (Figure 4, B and D). Growth rate and OD48 were significantly correlated with each other when BMN lines were grown in nystatin, consistent with the idea that both assays measure an aspect of growth rate (cor = 0.87, $t_{33} = 10.3$, $P < 0.0001$), and both were significantly correlated with nystatin IC_{50} (OD48, cor = 0.74, $t_{33} = 6.4$, $P < 0.0001$; growth rate, cor = 0.76, $t_{33} = 6.8$, $P < 0.0001$). When lines were grown in YPD, however, growth rate and OD48 were not significantly correlated with each other, consistent with growth having completed by 48 hr in YPD and OD48 measuring solely the efficiency of converting resources into cellular material (cor = 0.26, $t_{33} = 1.5$, $P = 0.13$). Growth rate in YPD was significantly correlated with IC_{50} in nystatin (cor = 0.56, $t_{33} = 2.8$, $P = 0.01$), while OD48 was not (cor = 0.04, $t_{33} = 0.18$, $P = 0.20$). Interestingly, the correlation between growth rate in YPD and IC_{50} in nystatin was positive; that is, mutations with the broadest tolerance to nystatin were among the best to grow in YPD. The ergosterol

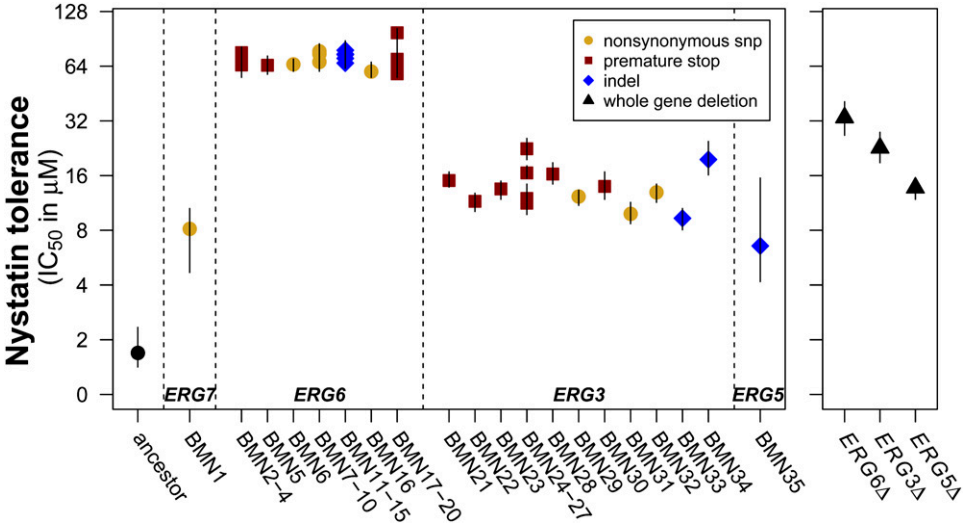


Figure 3 BMN lines have significantly increased tolerance to nystatin relative to the ancestor. Error bars represent 95% confidence intervals from likelihood profile plots. Lines grouped on the *x*-axis carried identical ergosterol mutations at the sequence level (Table 1). Mutation lines are arranged on the *x*-axis in the same way they were numbered, *i.e.*, based on which gene carries a mutation (mutations in genes farther from producing ergosterol are numbered lower and plotted to the left) and position in the gene (mutations closer to the start codon are numbered lower and plotted to the left of mutations closer to the stop codon).

gene that bore a mutation was significantly associated with both fitness proxies in both environments (growth rate in nystatin, $F_3 = 26.4$, $P < 0.0001$; OD48 in nystatin, $F_3 = 15.5$, $P < 0.0001$; growth rate in YPD, $F_3 = 25.8$, $P < 0.0001$; OD48 in YPD, $F_3 = 4.9$, $P = 0.007$). The type of mutation was found to have a significant effect on growth rate in nystatin ($F_2 = 5.0$, $P = 0.014$), although we note that of all significant statistical results this is the only one that does not remain significant when we combine multiple lines with the same ergosterol mutation (see File S1). In all other comparisons we found no significant association with the type of mutation (OD48 in nystatin, $F_2 = 0.86$, $P = 0.44$; growth rate in YPD, $F_2 = 0.27$, $P = 0.77$; OD48 in YPD, $F_2 = 0.09$, $P = 0.91$) nor was the interaction between ergosterol gene and type of mutation significant (growth rate in nystatin, $F_2 = 0.16$, $P = 0.85$; OD48 in nystatin, $F_2 = 0.68$, $P = 0.51$; growth rate in YPD, $F_2 = 0.33$, $P = 0.72$; OD48 in YPD, $F_2 = 0.73$, $P = 0.50$).

We observed substantial differences among the nystatin resistance lines in their tolerance to other stressful environments (breadth of tolerance measured as IC_{50} in all environments, Figure 5). We found significant negative correlations between tolerance to nystatin and tolerance to both copper and ethanol (copper, cor $= -0.80$, $t_{34} = -7.6$, $P < 0.00001$; ethanol, cor $= -0.63$, $t_{34} = -4.7$, $P < 0.00001$) and no correlation between nystatin and salt tolerance (cor $= -0.07$, $t_{34} = -0.39$, $P = 0.70$). The tolerance breadths exhibited by lines with mutations in the same ergosterol gene were fairly consistent, with only a few exceptions. Importantly, although reduced tolerance to all other environments tested was observed for some genes bearing nystatin resistance mutations (especially *ERG6* mutations), mutations in other ergosterol genes had no effect or even a positive effect on growth in the face of other environmental challenges (*e.g.*, positive fitness effects were observed for BMN-*erg7*, BMN-*erg3*, and BMN-*erg5* lines in copper). That is, mutations in different ergosterol genes exhibited significant sign $G \times E$ when comparing growth in nystatin and copper. Overall, the majority of lines differed in fitness from the ancestor in most environments (Figure 5, Table S8), but the pattern was heavily dependent on both environment and gene.

Discussion

Genes that act late in the ergosterol biosysnthesis pathway were found to be the primary (and possibly exclusive) target for the first step of adaptation by *S. cerevisiae* to low levels of nystatin. The mutational neighborhood was reasonably large, as we uncovered 20 unique mutations within four genes exhibiting increased tolerance to nystatin. At the gene level, however, the genomic scope for beneficial mutations was quite narrow in this environment, as all but two lines carried mutations within *ERG6* and *ERG3*. Lines with different mutations in the same gene tended to exhibit similar tolerance phenotypes in all environments tested, including altered levels of the original stressor, an unstressful environment (YPD), and three different stressful environments (copper, ethanol, and salt). We found nonparallel fitness effects of mutations in different ergosterol genes in the face of different environmental challenges, indicating the unpredictable nature of gene–environment interactions. Although some lines showed a decreased fitness in all other stressful environments tested (*i.e.*, BMN-*erg6* lines), other lines showed a mixture of fitness costs and benefits in other environments besides nystatin, with some lines having high tolerance in all environments tested (*i.e.*, BMN-*erg5*).

Parallel evolution is more likely to occur via loss-of-function mutations than via gain-of-function (Christin *et al.* 2010), and the different mutations that we observed in *ERG6*, *ERG3*, and *ERG5* may well have caused loss-of-function alleles. Consistent with this hypothesis, knockout lines for these three genes (*erg6*Δ, *erg3*Δ, and *erg5*Δ) are viable and have been shown to increase fitness in low levels of nystatin in a screen of all deletion collection lines (Hillenmeyer *et al.* 2008). Our own nystatin tolerance assay of these null

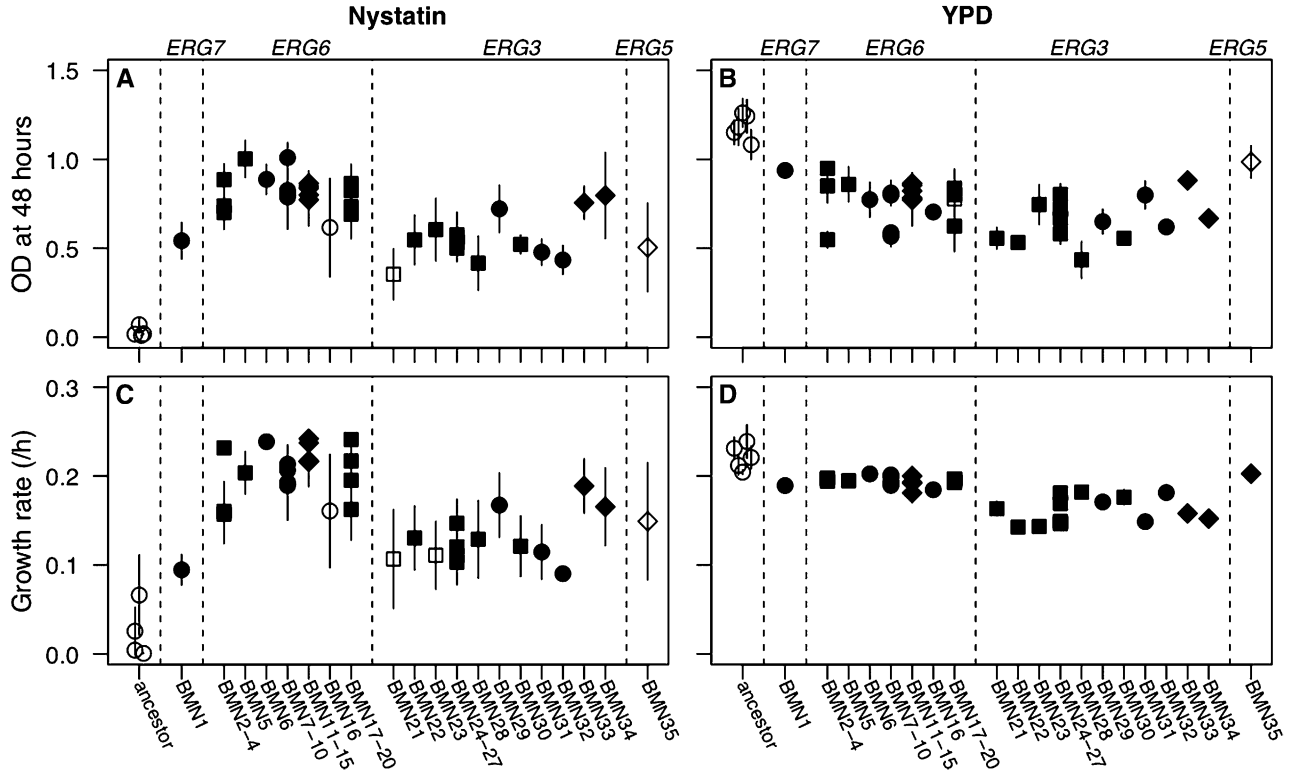


Figure 4 Two different fitness assays, OD at 48 hr (A, in nystatin; B, in YPD, an unstressful environment) and maximum growth rate (C, in nystatin; D, in YPD), show that BMN lines have increased growth in YPD + 4 μM nystatin and reduced growth in YPD relative to the ancestor. Lines that are significantly different from the ancestral colonies are plotted with solid symbols (*t*-test results presented in Table S4, Table S5, Table S6, and Table S7). Error bars depict the standard error of four bioscreen well replicates for each line. Mutation lines are arranged on the *x*-axis as in Figure 3. For the ancestral lines, five independent colonies were isolated and used to establish the lines, each of which was assayed four times (bars again show SE based on the replicates for each of these five lines). Mutations in the same gene are grouped together, and genes that are closer to producing ergosterol (the end product of the common pathway) are farther to the right. When multiple mutations are present in the same gene, mutations closer to the start codon are numbered lower.

mutations in nystatin showed subtle but significant differences between BMN-*erg6* and BMN-*erg3* lines and the appropriate knockout lines (Figure 3). These results suggest that our mutations are similar, but not identical, to the null mutations and may indicate that the enzymes these genes encode retain some activity. Different amino acid changes in *ERG6* have previously been shown to have different kinetic properties (Nes *et al.* 2004), and so we wish to be cautious in concluding that all of these mutations represent complete loss of function. By contrast to *ERG6* and *ERG3*, we identified only one mutation in *ERG5*. The lack of parallel mutations in *ERG5* is surprising, given that this is a longer gene (1616 bp) than either *ERG6* (1151 bp) or *ERG3* (1097 bp). Because *erg5*Δ is respiratory deficient (Merz and Westermann 2009), we checked 25 of the petite lines we isolated from our screens for sequence changes in *ERG5*, yet found no evidence of additional mutations in *ERG5*. The genomic DNA from four additional petite lines could not be amplified using two different sets of primers, and thus additional mutations in *ERG5* may be present in those lines. The fact that BMN35 was not respiratory deficient and could grow in glycerol (unlike *erg5*Δ) indicates that this mutant is not equivalent to a whole-gene knockout. We thus conjecture that relatively few adaptive mutations were found in *ERG5* because relatively rare changes were required to allow adaptation to nystatin without full loss of function.

BMN1 (with a mutation in *ERG7*) cannot be a loss-of-function mutation as *erg7*Δ is inviable. Consistent with this claim, the only mutation that arose in *ERG7* was a nonsynonymous change very close to the end of the gene. It is thus plausible that this particular change was a gain-of-function mutation. Furthermore, our sterol profile of BMN1 is unique among the mutations we acquired. Similarly, BMN1 does not seem to share a sterol profile with any of the mutants identified in earlier studies on nystatin resistance, although sterol profiles in these early mutants match the profiles exhibited by our BMN lines that carry mutations in *ERG6*, *ERG3*, and *ERG5* (Woods 1971; Bard 1972; Grunwald-raij and Margalith 1990). The unique phenotype generated by the mutation in *ERG7* deserves future investigation. It may be that other gain-of-function mutations could have been beneficial in nystatin but were not sampled due to rarity or a bias toward large-effect loss-of-function mutations in our screens. Interestingly, *S. cerevisiae* knockout lines for all other genes that act late in the ergosterol pathway

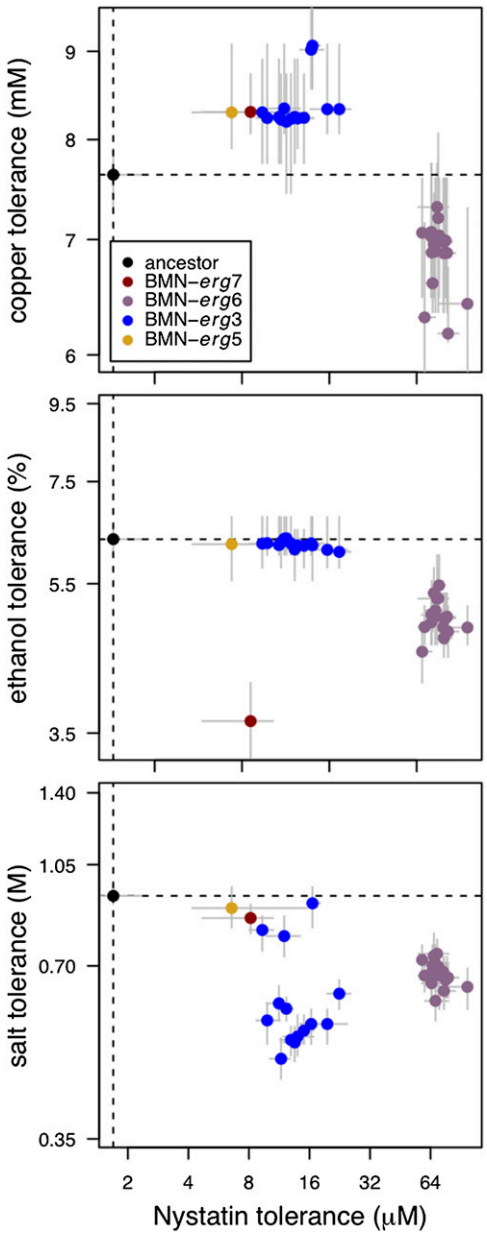


Figure 5 Tolerance to nystatin does not predict tolerance to other stressful environments. Tolerance (measured as IC_{50} in all environments) was measured in copper ($CuSO_4$), ethanol, and salt (NaCl). Dashed lines indicate the ancestral tolerance in each environment. Lines that appeared to have very similar tolerance in one environment did not necessarily have a similar tolerance in a second one, indicating significant $G \times E$ interactions.

(except *hmg1*Δ and *hmg2*Δ, which are isozymes, so that deleting either one alone is not expected to have a strong effect on growth) are inviable [*erg10*Δ, *erg13*Δ, *erg12*Δ, *erg8*Δ, *mvd1*Δ, *erg20*Δ, *erg9*Δ, *erg1*Δ, *erg11*Δ, and *erg25*Δ (Giaver *et al.* 2002)], are unable to grow aerobically under our growth conditions [*erg24*Δ (Lees *et al.* 1995)], are ergosterol auxotrophs [*erg2*Δ (Parks and Casey 1995)], or have reduced fitness in nystatin [*erg4*Δ (Hillenmeyer *et al.* 2008)], helping to explain the narrow gene target of adaptation to nystatin that we have observed.

The distribution of fitness effects of beneficial mutations is an important factor that dictates how populations might adapt to a novel stressor. In his seminal work on adaptive mutations, Gillespie used the extreme value theory to suggest that one-step beneficial mutations might be expected to exhibit exponentially distributed selective effects (Gillespie 1983, 1984, 1991). The 20 unique one-step nystatin adaptive mutations we have acquired here do not immediately appear to fit this prediction, as we recovered an abundance of large-effect mutations (Figure 3) whose tolerance to nystatin far exceeds the 4 μM exposure concentration at which they were acquired. A number of explanations contribute to this finding. Our assay would not have detected small-effect mutations, because we required mutations of large enough effect to enable growth in 4 μM nystatin. Furthermore, the mutations are not independent, as many are in the same genes. As discussed above, we suspect that the 7 different mutations in *ERG6* and 11 unique mutations in *ERG3* are largely loss-of-function mutations in the ergosterol pathway. In any environment where large-effect loss-of-function mutations are available, many different nonsynonymous SNPs or indels could be selected, and we might expect these mutations to skew the distribution of beneficial effects toward the maximal fitness effect possible via eliminating the target of selection, here ergosterol. We also expect that the first mutations acquired have a distribution skewed toward large-effect mutations, compared to the distribution of all possible beneficial mutations, because of their selective advantage. Our results are thus more consistent with the extreme value properties in the Weibull domain (where there is a maximal fitness benefit) than in the Gumbel domain used by Gillespie (Joyce *et al.* 2008).

We found that tolerance across environments frequently exhibited gene–environment interactions, which were typically consistent across different mutations within the same gene. All BMN lines had a decreased growth rate and decreased biomass production (OD48) in the unstressful environment, YPD. Interestingly, we found no evidence that mutations with a larger benefit in nystatin had a greater negative effect in other environments. This is highlighted by a significant positive correlation between IC_{50} in nystatin and growth rate in YPD. When we examined growth in three other stressful environments (copper, ethanol, and salt), we found that beneficial mutations to nystatin had pleiotropic effects that differed substantially among environments ($G \times E$). For example, while all ergosterol mutations examined were beneficial in nystatin, *ERG6* and *ERG7* mutations had reduced tolerance to ethanol, while *ERG3* and *ERG5* mutations were very similar in tolerance to the ancestor. The $G \times E$ interactions were so extreme that some mutations exhibited opposite selective effects in some environments. In particular, *ERG6* mutations were less tolerant to copper, while *ERG3*, *ERG5*, and *ERG7* were more tolerant. We call this phenomenon, where two mutations that are beneficial in one environment have selective effects that differ in sign in another environment, sign $G \times E$

[by analogy to "sign epistasis" (Weinreich *et al.* 2005)]. These experiments demonstrate that although adaptive mutations may show parallel phenotypes in a particular environment (here, in nystatin), effects in other environments of interest can be idiosyncratic and must be specifically examined.

Our results demonstrate that even with a narrow target for adaptation at the gene level (all 35 lines isolated in our screens carried mutations in only four different genes), mutations that appear phenotypically similar in one environment may well present variability in others. As a final example from our data set, although BMN35 with a mutation in *ERG5* shares a similar IC_{50} phenotype with BMN-*erg3* lines in nystatin, ethanol, and copper, it has a very different phenotype in salt. Consequently, different subsets of adaptive mutations are likely to be favorable under environmental conditions that require adaptation to more than one selective agent. Furthermore, our results demonstrate that if different first mutations are acquired by separated populations during adaptation to nystatin, this may well place different populations at different locations on the adaptive landscape following shifts in other environmental variables, altering the future evolutionary pathways accessible to these populations. The ability to sequence the entire genomes of multiple adapting lines provides an extremely useful way to explore the range of genetic pathways that evolution can take.

Acknowledgments

The authors thank R. FitzJohn for assistance with R; A. Kuzmin for Illumina library construction; N. Kane for many helpful discussions on the analysis of next generation sequence data; S. Lee, W. Li, A. Van Tol, M. Campbell, and J. Ono for laboratory assistance; and the Otto Laboratory group, F. DéBarre, J. Hill, and two anonymous reviewers for helpful comments on the manuscript. This work was supported by the Canadian National Science and Engineering Research Council (A.C.G. and S.P.O.) and the Killam Trusts (A.C.G.).

Literature Cited

Anderson, J. B., C. Sirjusingh, A. B. Parsons, C. Boone, C. Wickens *et al.*, 2003 Mode of selection and experimental evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. Genetics 163: 1287–1298.

Araya, C. L., C. Payen, M. J. Dunham, and S. Fields, 2010 Whole-genome sequencing of a laboratory-evolved yeast strain. BMC Genomics 11: 88.

Arthington-Skaggs, B. A., H. Jradi, T. Desai, and C. J. Morrison, 1999 Quantitation of ergosterol content: novel method for determination of fluconazole susceptibility of *Candida albicans*. J. Clin. Microbiol. 37: 3332–3337.

Bard, M., 1972 Biochemical and genetic aspects of nystatin resistance in *Saccharomyces cerevisiae*. J. Bacteriol. 111: 649–657.

Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh *et al.*, 2009 Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature 461: 1243–1247.

Bataillon, T., T. Zhang, and R. Kassen, 2011 Cost of adaptation and fitness effects of beneficial mutations in *Pseudomonas fluorescens*. Genetics 189: 939–949.

Bennett, A., and R. Lenski, 2007 An experimental test of evolutionary trade-offs during temperature adaptation. Proc. Natl. Acad. Sci. USA 104: 8649.

Bhiyan, M. S. A., Y. Ito, A. Nakamura, N. Tanaka, K. Fujita *et al.*, 1999 Nystatin effects on vacuolar function in *Saccharomyces cerevisiae*. Bioscience 63: 1075–1082.

Burch, C., and L. Chao, 2000 Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature 406: 625–628.

Carrillo-Munoz, A. J., G. Giusiano, P. A. Ezkurra, and G. Quindós, 2006 Antifungal agents: mode of action in yeast cells. Rev. Esp. Quimioter. 19: 130–139.

Chou, H. H., J. Berthet, and C. J. Marx, 2009 Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. PLoS Genet. 5: e1000652.

Christin, P.-A., D. M. Weinreich, and G. Besnard, 2010 Causes and evolutionary significance of genetic convergence. Trends Genet. 26: 400–405.

Conrad, T. M., N. E. Lewis, and B. Palsson, 2011 Microbial laboratory evolution in the era of genome-scale science. Mol. Syst. Biol. 7: 1–11.

Cooper, T. F., D. E. Rozen, and R. E. Lenski, 2003 Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 100: 1072–1077.

FitzJohn, R. G., W. P. Maddison, and S. P. Otto, 2009 Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Syst. Biol. 58: 595–611.

Gerstein, A. C., and S. P. Otto, 2011 Cryptic fitness advantage: diploids invade haploid populations despite lacking any apparent advantage as measured by standard fitness assays. PLoS ONE 6: e26599.

Ghannoum, M. A., and L. B. Rice, 1999 Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. Clin. Microbiol. Rev. 12: 501–517.

Giaver, G., A. M. Chu, L. Ni, and C. Connelly, L. Riles *et al.*, 2002 Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418: 387–391.

Gillespie, J., 1991 *The Causes of Molecular Evolution*. Oxford University Press, New York.

Gillespie, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. 23: 202–215.

Gillespie, J. H., 1984 Molecular evolution over the mutational landscape. Evolution 38: 1116–1129.

Gresham, D., M. Desai, C. Tucker, H. Jenq, D. Pai *et al.*, 2008 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet. 4: e1000303.

Grunwald-raij, H., and P. Margalith, 1990 Ethanol fermentation by nystatin-resistant strains of *Saccharomyces cerevisiae*. J. Appl. Microbiol. 68: 247–252.

Hapala, I., V. Klobucnikova, K. Mazanova, and P. Kohut, 2005 Two mutants selectively resistant to polyenes reveal distinct mechanisms of antifungal activity by nystatin and amphotericin b. Biochem. Soc. Trans. 33: 1206–1209.

Hereford, J., 2009 A quantitative survey of local adaptation and fitness trade-offs. Am. Nat. 173: 579–588.

Herring, C. D., A. Raghunathan, C. Honisch, T. Patel, M. K. Applebee *et al.*, 2006 Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. 38: 1406–1412.

Hillenmeyer, M. E., E. Fung, J. Wildenhain, S. E. Pierce, S. Hoon *et al.*, 2008 The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science 320: 362–365.

Jensen-Pergakes, K. L., M. A. Kennedy, N. D. Lees, R. Barbuch, C. Koegel *et al.*, 1998 Sequencing, disruption, and characterization of the *Candida albicans* sterol methyltransferase (*ERG6*) gene: drug susceptibility studies in *erg6* mutants. Antimicrob. Agents Chemother. 42: 1160–1167.

Joyce, P., D. R. Rokyta, C. J. Beisel, and H. A. Orr, 2008 A general extreme value theory model for the adaptation of DNA sequences under strong selection and weak mutation. Genetics 180: 1627–1643.

Kanafani, Z. A., and J. R. Perfect, 2008 Resistance to antifungal agents: mechanisms and clinical impact. Clin. Infect. Dis. 46: 120–128.

Kishimoto, T., L. Iijima, M. Tatsumi, N. Ono, A. Oyake *et al.*, 2010 Transition from positive to neutral in mutation fixation along with continuing rising fitness in thermal adaptive evolution. PLoS Genet. 6: e1001164.

Lees, N. D., B. Skaggs, D. R. Kirsch, and M. Bard, 1995 Cloning of the late genes in the ergosterol biosynthetic pathway of *Saccharomyces cerevisiae* – a review. Lipids 30: 221–226.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and samtools. Bioinformatics 25: 2078–2079.

Lynch, M., W. Sung, K. Morris, N. Coffey, and C. Landry, 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc. Natl. Acad. Sci. USA 105: 9272–9277.

MacLean, R. C., and A. Buckling, 2009 The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. PLoS Genet. 5: e1000406.

Merz, S., and B. Westermann, 2009 Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in *Saccharomyces cerevisiae*. Genome Biol. 10: R95.

Minty, J. J., A. A. Lesnefsky, F. Lin, Y. Chen, T. A. Zaroff *et al.*, 2011 Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in *Escherichia coli*. Microb. Cell Fact. 10: 18.

Mukhopadhyay, K., A. Kohli, and R. Prasad, 2002 Drug susceptibilities of yeast cells are affected by membrane lipid composition. Antimicrob. Agents Chemother. 46: 3695–3705.

Nes, W., P. Jayasimha, W. Zhou, R. Kanagasabai, C. Jin *et al.*, 2004 Sterol methyltransferase: functional analysis of highly conserved residues by site-directed mutagenesis. Biochemistry 43: 569–576.

Ostrowski, E., D. Rozen, and R. Lenski, 2005 Pleiotropic effects of beneficial mutations in *Escherichia coli*. Evolution 59: 2343–2352.

Ostrowski, E., R. Woods, and R. Lenski, 2008 The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. Proc. Biol. Sci. 275: 277–284.

Parks, L. W., and W. M. Casey, 1995 Physiological implications of sterol biosynthesis in yeast. Annu. Rev. Microbiol. 49: 95–116.

Pelosi, L., L. Kühn, D. Guetta, J. Garin, J. Geiselmann *et al.*, 2006 Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. Genetics 173: 1851–1869.

Poole, A. M., M. J. Phillips, and D. Penny, 2003 Prokaryote and eukaryote evolvability. Biosystems 69: 163–185.

Pörtner, H., A. Bennett, F. Bozinovic, A. Clarke, M. Lardies *et al.*, 2006 Trade-offs in thermal adaptation: the need for a molecular to ecological integration. Physiol. Biochem. Zool. 79: 295–313.

R Development Core Team, 2011 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Rokyta, D. R., P. Joyce, S. B. Caudle, and H. A. Wichman, 2005 An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. Nat. Genet. 37: 441–444.

Sambrook, J., and D. W. Russell, 2001 *Molecular Cloning: A Laboratory Manual*, Ed. 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schluter, D., 2009 Evidence for ecological speciation and its alternative. Science 323: 737.

Selmecki, A. M., K. Dulmage, L. E. Cowen, J. B. Anderson, and J. Berman, 2009 Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet. 5: e1000705.

Tenaillon, O., A. Rodriguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett *et al.*, 2012 The molecular diversity of adaptive convergence. Science 335: 457–461.

Toprak, E., A. Veres, J.-B. Michel, R. Chait, D. L. Hartl *et al.*, 2011 Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. 44: 101–105.

Weinreich, D., R. Watson, and L. Chao, 2005 Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59: 1165–1174.

Wong, A., and R. Kassen, 2011 Parallel evolution and local differentiation in quinolone resistance in *Pseudomonas aeruginosa*. Microbiology 157: 937–944.

Woods, R., 1971 Nystatin-resistant mutants of yeast: alterations in sterol content. J. Bacteriol. 108: 69–73.

Woods, R., D. Schneider, C. L. Winkworth, M. A. Riley, and R. E. Lenski, 2006 Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc. Natl. Acad. Sci. USA 103: 9107–9112.

Communicating editor: J. J. Bull

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2012/06/19/genetics.112.142620.DC1

Parallel Genetic Changes and Nonparallel Gene–Environment Interactions Characterize the Evolution of Drug Resistance in Yeast

Aleeza C. Gerstein, Dara S. Lo, and Sarah P. Otto

DOI: 10.1534/genetics.112.142620

File S1

APPEARANCE OF IDENTICAL MUTATIONS

Five identical ergosterol mutations were sampled within multiple lines (Table 1). The most likely explanation is that these mutations arose during population expansion before the lineages were isolated from one another and before the stressor was applied. The culture used to seed mutation acquisition screens was derived from a single wild type colony grown up overnight in YPD, an unstressful environment. Given the size of this overnight culture ($\sim 1.2 \times 10^9$ cells), there would have been approximately 30 generations of growth during this YPD phase ($2^{30} = 1.1 \times 10^9$). Despite the bottleneck to a single colony-forming unit, the population size of the source population, a plausible per-base pair mutation rate (0.33×10^{-9}, LYNCH *et al.* 2008), the hundreds of one-step mutations that could potentially result in nystatin tolerance ($\sim$ 350 different mutations based on our data of what types of mutations confer tolerance to nystatin, see section below), and the number of founding lineages (60 and 180 in screens 'a' and 'b', respectively), we calculated that there is a high probability that some beneficial mutations were segregating in the precursor population (see File S2). Standing genetic variation from a colony grown for a single overnight in YPD has previously been found to play a large role in the eventual mutations that were selected in a yeast experimental evolution project (GRESHAM *et al.* 2008), and we believe this is also the most likely explanation here.

Well-to-well contamination is also possible, yet unlikely. We kept track of where mutations were isolated within the 96 well plates; in no case was the same mutation isolated in neighbouring wells and in multiple cases the same mutation was isolated from different plates within the same screen. In at least the case of BMN11-15, well-to-well contamination is even less likely, as growth was seen in all wells before the first culture was isolated, thus there was little to no opportunity for contamination from one well to another (Table S1).

Although certainly possible biologically, we do not have any support for the same allelic variant arising independently in multiple lines. Mutation lines were acquired in two separate screens (denoted 'a' and 'b' in Table S1). The exact same protocol was used for both acquisition screens (see Methods), however, a different colony was grown up overnight to provide the culture used to seed all replicate wells in each screen. In no case was the same allele sampled in the two different screens.

EFFECT OF NON-ERGOSTEROL MUTATIONS

While we attempted to minimize the number of mutations carried by the lines, secondary mutations could have arisen and fixed during the $\sim$30 generations of growth in YPD within the precursor population or during the $\sim$30 generations of growth in nystatin required for yeast precipitate to be detected. Nine lines carry non-synonymous nuclear mutations in genes not involved in the ergosterol pathway (Table S2). Two of these lines share the same ergosterol mutation

with other BMN lines, which allows us to directly assess the phenotypic affect of these additional mutations. BMN27 carries three additional mutations: nonsynonymous mutations in YJR107W (an uncharacterized protein), *AUR1* (a protein required for sphingolipid synthesis), and an extra copy of chromosome 2. BMN27 has a higher IC_{50} in both salt and copper than the three other lines that carry the same ergosterol mutation (BMN24-26), but these differences are not significant. BMN15, with a nonsynonymous mutation in *MBP1*, also does not differ in our fitness assays from BMN11-14, lines with which it shares *ERG6* and *GDA1* mutations. The remaining lines with secondary mutations have very similar nystatin tolerance to other lines that carry mutations in the same ergosterol gene. As a further test, we backcrossed and sporulated representative lines that contain a mutation in each of the four ergosterol genes (BMN1, BMN9, BMN32 and BMN58) to BY4739. For each backcross we found 2:2 segregation of nystatin tolerance. We thus have little reason to suspect that mutations in non-ergosterol genes are strongly influencing our results.

STATISTICAL RESULTS REMAIN THE SAME IF WE COMBINE LINES WITH THE SAME ERGOSTEROL MUTATION

Statistical results reported in the main text are upheld if we use the average tolerance and fitness results from lines that contain the same ergosterol mutation. For all three assays conducted in nystatin (IC_{50}, growth rate, $OD48$), and both assays conducted in YPD (growth rate and $OD48$), we recover the same results previously reported, only the ergosterol gene that bears a mutation has a significant effect on the results (i.e., mutation type and their interaction do not).

The statistical results of our correlation tests between different fitness proxies also yield the same results if we average across line replicates with the same mutation. Comparing between nystatin tolerance breadth (IC_{50}) and two fitness proxies in nystatin we find that all three assays are significantly correlated to each other (growth rate and $OD48$: cor = 0.88, t_{18} = 7.9, $p < 0.0001$; growth rate and IC_{50}: cor = 0.80, t_{18} = 5.7, $p < 0.0001$; $OD48$ and IC_{50}: cor = 0.72, t_{18}= 4.4,p = 0.0003). When we compare IC_{50} in nystatin and the same two fitness proxies when the lines are grown in YPD we find the same result we previously reported, i.e., the only significant correlation is growth rates in YPD with IC_{50} in nystatin (growth rate and $OD48$ in YPD: cor = 0.35, t_{18} = 1.6, p = 0.13; growth rate in YPD and IC_{50} in nystatin: cor = 0.65, t_{18} = 3.6, p = 0.002; $OD48$ and IC_{50} in nystatin: cor = 0.20, t_{18} = 0.9, p = 0.4).

Finally, we also recover the same pattern of tradeoffs between tolerance to nystatin and secondary environments (all measured as IC_{50}). Specifically, we find tolerance to nystatin is significantly correlated to both ethanol (cor = -0.62, t_{19} = -3.5, p = 0.002) and copper (cor = -0.88, t_{19} = -8.0, $p < 0.0001$), but not to salt (cor = 0.13, t_{19} = 0.6, p = 0.56).

LITERATURE CITED

GRESHAM, D., M. DESAI, C. TUCKER, H. JENQ, D. PAI, *et al.*, 2008 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet **4**: e1000303.

LYNCH, M., W. SUNG, K. MORRIS, N. COFFEY, and C. LANDRY, 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. PNAS **105**: 9272–9277.

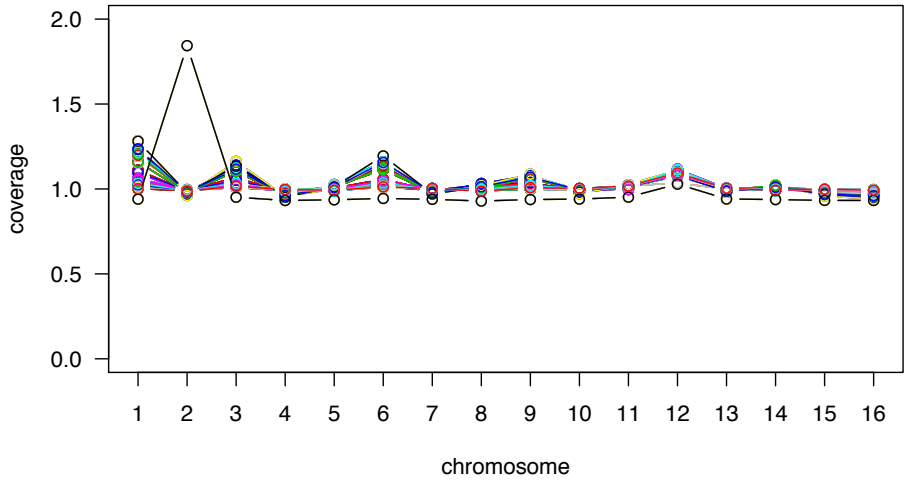


Figure S1 Relative coverage of each chromosome from genomic alignments. Using the Illumina genomic sequence data, the total coverage for each chromosome was calculated as the proportion of sequenced sites mapping to a particular chromosome relative to the proportion of known mapped sites located on that chromosome within the yeast reference genome (as reported by configureBuild.pl in Illumina's CASAVA-1.8.0 package). Examining the coverage data for each chromosome from each BMN line (each line is plotted with a unique colour) indicates only one aneuploidy event - an extra copy of chromosome 2 in BMN27.

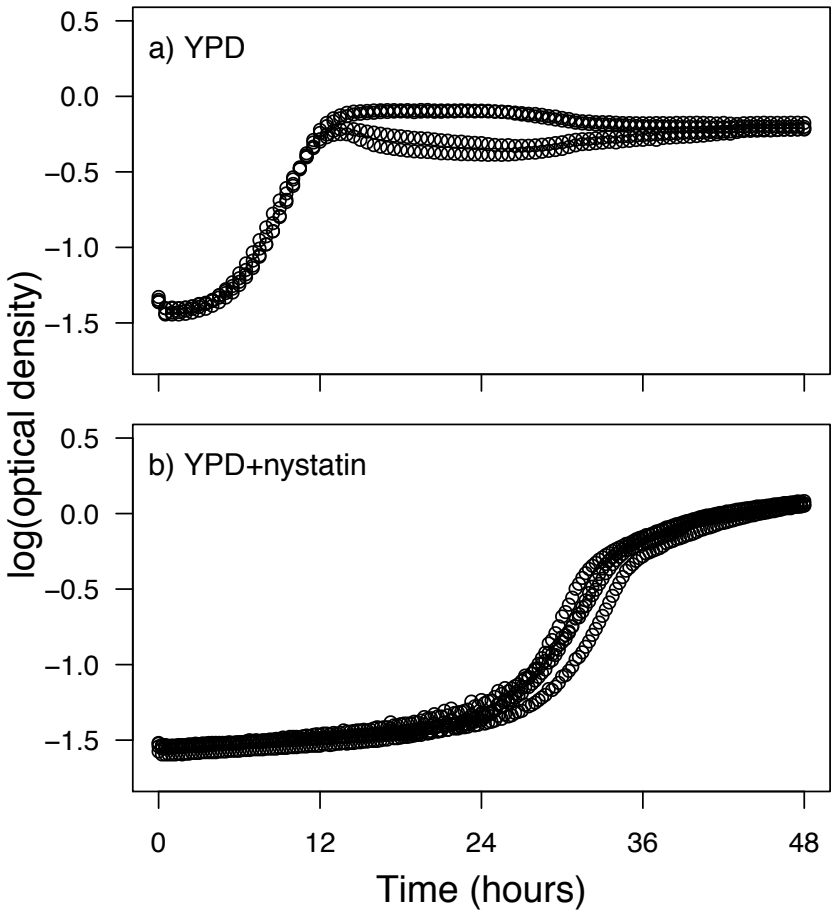


Figure S2 Representative raw growth curves for one line (BMN31) in a) YPD and b) YPD+4μM nystatin. Optical density was measured automatically by a Bioscreen C Microbiology Workstation. Optical density has stopped increasing by 48 hours in YPD, while lines are still growing at 48 hours in nystatin. The optical density at 48 hours ($OD48$) thus reflects primarily resource efficiency in YPD while it represents a combined measure of resource efficiency and growth rate in nystatin.

Table S1 The date mutations were acquired. Mutations were acquired in two screens ('a' & 'b'), with each acquisition experiment lasting one week. Different ancestral colonies were used to initiate each acquisition screen. By examining the date of isolation and screen we gain insight into the process that led to identical ergosterol mutations in multiple lines (BMN2-4, BMN7-10, BMN11-15, BMN17-20 and BMN24-27).

BMN	Date Isolated	Screen
1	10.07.25	a
2	10.08.14	b
3	10.08.14	b
4	10.08.09	b
5	10.07.23	a
6	10.07.25	a
7	10.07.25	a
8	10.07.25	a
9	10.07.25	a
10	10.07.23	a
11	10.08.09	b
12	10.08.09	b
13	10.08.09	b
14	10.08.09	b
15	10.08.09	a
16	10.07.23	a
17	10.07.25	a
18	10.07.25	a
19	10.07.23	a
20	10.07.23	a
21	10.08.10	b
22	10.08.12	b
23	10.08.12	b
24	10.08.10	b
25	10.08.12	b
26	10.08.12	b
27	10.08.13	b
28	10.07.25	a
29	10.07.25	a
30	10.08.10	b
31	10.08.10	b
32	10.08.12	b
33	10.08.10	b
34	10.08.13	b
35	10.08.14	b

Table S2 Genotypic basis of mutations in genes not in the ergosterol biosynthesis pathway. Nineteen additional unique mutations plus one aneuploidy were found within 19 lines.

BMN Line	Gene	Genome Position (Chr.Bp)	Gene Position (in nucleotides)	Mutation	Amino Acid Change
1	YOL073C	XV.193885	916	G>A	Asp306Asn
3	YPL039W	XVI.479629	405	G>A	synonymous (Arg)
5	*CDC23*	VIII.438829	222	A>C	synonymous (Ile)
8		X.30640		A>C	
11-15	*GDA1*	V.74568	798	C>T	synonymous (Val)
15	*MBP1*	IV.354071	1195	T>G	Phe399Val
16		XIV.1753512		A>T	
16		XIV.1753521		C>A	
22	*COX1*	mt.23360	9543	T>A	synonymous (Ile)
22	*COX1*	mt.23361	9544	A>T	Ile3182Leu
23	*SCW11*	VII.442319	591	A>G	synonymous (Ser)
24-27	*FCY2*	V.267873	241	G>A	Glu81Lys
26		XIV.507563		T>G	
27	*AUR1*	XI.436609	1030	C>T	Pro344Ser
27	YJR107W	X.627995	656	G>A	Trp219Stop
27				+Chr2	
29	*SGS1*	XIII.644130	1129	A>G	Asn377Asp
30	*MDM20*	XV.188973	1950	T>G	Ile650Met
31	*ALT1*	XII.319765	251	T>C	Leu84Pro
35		I.73925		A>G	

Table S3 Maximum likelihood results for growth in nystatin. We fit a likelihood model to the combined data for each mutation line and the ancestral line. The full model allowed two IC_{50} values to be fit to the data, while the constrained model forced the IC_{50} for the mutant and ancestral lines to be equal. All mutation lines were found to have significantly different IC_{50} values from the ancestor (difference in log-likelihood between the two models > 1.92; see Methods).

BMN	Gene	LogLikelihood, FullModel	LogLikelihood, Constrained	Difference in Models
	*erg6*Δ	35.60	-12.04	47.64
	*erg5*Δ	18.18	-10.64	28.82
	*erg3*Δ	46.23	5.76	40.46
1	*ERG7*	18.12	10.77	7.34
2	*ERG6*	50.14	-5.47	55.61
3	*ERG6*	38.97	-16.94	55.91
4	*ERG6*	60.91	-18.95	79.86
5	*ERG6*	37.31	-13.68	50.99
6	*ERG6*	60.87	-0.47	61.34
7	*ERG6*	53.25	-2.56	55.81
8	*ERG6*	49.38	-16.30	65.68
9	*ERG6*	35.15	-16.92	52.08
10	*ERG6*	42.37	-13.91	56.28
11	*ERG6*	62.46	6.72	55.74
12	*ERG6*	63.21	7.17	56.04
13	*ERG6*	69.14	12.87	56.27
14	*ERG6*	38.78	-16.81	55.59
15	*ERG6*	51.53	-17.35	68.89
16	*ERG6*	39.39	-15.73	55.12
17	*ERG6*	68.80	12.80	56.00
18	*ERG6*	35.34	-13.35	48.69
19	*ERG6*	35.71	-16.51	52.22
20	*ERG6*	42.61	-12.49	55.10
21	*ERG3*	43.94	8.56	35.37
22	*ERG3*	33.01	-0.17	33.18
23	*ERG3*	31.41	-0.68	32.09
24	*ERG3*	31.01	6.83	24.18
25	*ERG3*	28.70	2.33	26.37
26	*ERG3*	57.92	14.94	42.98
27	*ERG3*	60.53	23.06	37.47
28	*ERG3*	34.21	-3.92	38.13
29	*ERG3*	41.40	4.49	36.91
30	*ERG3*	48.29	17.20	31.09
31	*ERG3*	29.86	-0.50	30.36
32	*ERG3*	35.36	9.62	25.74
33	*ERG3*	29.21	-0.11	29.32
34	*ERG3*	25.92	1.46	24.46
35	*ERG5*	-3.30	-14.00	10.70

Table S4 T-test results comparing growth rate of BMN lines in nystatin to five ancestral colonies.

BMN	t	df	p-value
1	-3.4	18.7	0.003
2	-15.5	35.8	< 0.0001
3	-4.1	6.9	0.005
4	-4.1	6.8	0.005
5	-7.1	8.7	< 0.0001
6	-16	33.6	< 0.0001
7	-5.9	9.3	0.0002
8	-12.4	20.3	< 0.0001
9	-7.6	14.6	< 0.0001
10	-4.1	6.1	0.006
11	-16.8	33.6	< 0.0001
12	-7.6	8.6	< 0.0001
13	-9.1	7.1	< 0.0001
14	-15.8	30.8	< 0.0001
15	-6.6	7.5	0.0002
16	-2.4	3.3	0.087
17	-5.4	8.4	0.0005
18	-4.0	9.2	0.003
19	-6.7	7.6	0.0002
20	-17	32.5	< 0.0001
21	-1.7	3.4	0.18
22	-3.1	5.3	0.026
23	-2.4	5.2	0.06
24	-4.4	7.8	0.002
25	-3.1	10.3	0.011
26	-6.2	23.0	< 0.0001
27	-2.6	6.8	0.034
28	-3.5	3.7	0.027
29	-2.5	6.0	0.046
30	-4.1	6.5	0.0056
31	-2.9	5.5	0.029
32	-2.9	7.2	0.021
33	-5.1	33.8	< 0.0001
34	-5.4	7.2	0.0009
35	-2.1	4.4	0.10

Table S5 T-test results comparing $OD48$ of BMN lines grown in nystatin to five ancestral colonies.

BMN	t	df	*p*-value
1	-5.5	8.2	0.0005
2	-10.8	5.2	0.0001
3	-8.1	5.1	0.0004
4	-7.1	5.1	0.0008
5	-10.5	5.1	0.0001
6	-12.1	3.1	0.001
7	-9.2	6.2	0.0001
8	-13.7	3.1	0.0007
9	-8.7	8.2	< 0.0001
10	-4.5	5.0	0.0066
11	-9.8	3.1	0.002
12	-10.9	5.2	0.0001
13	-14.0	3.2	0.0006
14	-6.4	2.0	0.023
15	-8.4	5.1	0.0004
16	-2.6	3.0	0.082
17	-5.9	6.1	0.001
18	-5.3	7.1	0.001
19	-6.1	5.1	0.002
20	-11.3	4.1	0.0003
21	-2.8	3.0	0.065
22	-4.4	4.1	0.012
23	-3.8	4.0	0.018
24	-7.2	5.2	0.0007
25	-8.2	6.2	0.0001
26	-8.8	4.3	0.0007
27	-4.4	5.1	0.007
28	-3.8	3.0	0.032
29	-3.0	5.1	0.030
30	-6.0	5.1	0.002
31	-10.8	4.4	0.0003
32	-7.0	5.2	0.0007
33	-6.1	3.1	0.008
34	-8.8	5.1	0.0003
35	-2.3	4.0	0.087

Table S6 T-test results comparing growth rate of BMN lines grown in YPD to five ancestral colonies.

BMN	t	df	*p*-value
1	6.0	20.6	< 0.0001
2	4	20.1	0.0008
3	3.7	13.7	0.0026
4	4.7	21.0	0.0001
5	5	20.4	0.0001
6	3.3	21.6	0.0033
7	3.5	21.8	0.0019
8	3.2	7.0	0.0146
9	3.9	21.8	0.0007
10	5.8	21.8	< 0.0001
11	3.9	5.2	0.011
12	4.7	18.0	0.0002
13	3.7	9.1	0.0045
14	4.1	19.9	0.0006
15	4.9	20.4	0.0001
16	4.9	10.1	0.0006
17	3.9	17.0	0.0012
18	4.4	22.0	0.0002
19	5.0	20.1	0.0001
20	5.5	19.4	< 0.0001
21	6.8	7.2	0.0002
22	14.3	21.8	< 0.0001
23	14.0	22.0	< 0.0001
24	6.8	9.3	0.0001
25	10.9	12.6	< 0.0001
26	13.4	21.3	< 0.0001
27	5.6	11.1	0.0002
28	10.1	12.9	< 0.0001
29	5.1	8.8	0.0007
30	7.6	14.1	< 0.0001
31	5.1	6.6	0.0017
32	12.6	21.6	< 0.0001
33	4.5	6.7	0.0029
34	7.5	7.4	0.0001
35	2.6	11.4	0.022

Table S7 T-test results comparing $OD48$ of BMN lines grown in YPD to five ancestral colonies.

BMN	t	df	*p*-value
1	5.5	12.0	0.0001
2	6.0	19.0	<0.0001
3	3.8	4.0	0.020
4	12.2	7.9	< 0.0001
5	3.5	4.0	0.024
6	4.5	4.0	0.011
7	10.1	5.9	0.0001
8	7.3	7.7	0.0001
9	12.3	9.5	< 0.0001
10	5.3	4.9	0.003
11	5.2	5.9	0.0022
12	4.7	4.9	0.0058
13	7.0	8.2	0.0001
14	3.1	3.4	0.043
15	8.8	11.5	< 0.0001
16	13.7	21.8	< 0.0001
17	3.6	3.7	0.026
18	2.8	3.3	0.063
19	4.4	3.4	0.016
20	3.5	3.8	0.027
21	10.1	5.7	0.0001
22	17.8	21.7	< 0.0001
23	4.3	3.7	0.015
24	6.2	5.9	0.0009
25	8.0	5.5	0.0003
26	9.2	8.6	< 0.0001
27	10.0	6.0	0.0001
28	10.6	9.5	< 0.0001
29	7.8	3.9	0.0016
30	7.7	5.0	0.0006
31	18.4	20.4	< 0.0001
32	5.0	4.5	0.0052
33	15.7	22.0	< 0.0001
34	5.7	7.6	0.0006
35	2.3	4.1	0.079

Table S8 Likelihood ratio tests comparing IC_{50} of ancestral and BMN lines in copper, ethanol and salt.

		Copper (CuSO4)	Ethanol	Salt (NaCl)
BMN	Gene	Difference in Models	Difference in Models	Difference in Models
1	*ERG7*	3.76	30.00	4.79
2	*ERG6*	5.30	11.75	23.87
3	*ERG6*	0.62	5.18	23.50
4	*ERG6*	2.15	12.06	27.00
5	*ERG6*	1.73	24.35	26.36
6	*ERG6*	12.24	25.31	19.81
7	*ERG6*	1.93	14.42	20.05
8	*ERG6*	2.58	17.05	26.35
9	*ERG6*	2.61	17.76	30.26
10	*ERG6*	3.05	15.25	34.54
11	*ERG6*	2.99	8.05	18.08
12	*ERG6*	2.39	4.19	25.83
13	*ERG6*	1.88	18.78	32.90
14	*ERG6*	13.63	17.55	27.48
15	*ERG6*	1.91	29.10	29.00
16	*ERG6*	4.34	20.17	26.91
17	*ERG6*	0.44	6.95	23.83
18	*ERG6*	3.49	21.75	27.88
19	*ERG6*	1.21	20.40	24.13
20	*ERG6*	1.94	33.79	28.62
21	*ERG3*	2.20	0.47	48.55
22	*ERG3*	2.00	0.02	49.20
23	*ERG3*	2.50	1.00	48.42
24	*ERG3*	10.06	0.02	9.90
25	*ERG3*	2.17	0.15	40.52
26	*ERG3*	5.49	0.48	40.25
27	*ERG3*	9.99	0.49	0.44
28	*ERG3*	5.48	0.58	45.23
29	*ERG3*	9.82	< 0.0001	45.53
30	*ERG3*	1.18	0.01	44.65
31	*ERG3*	3.20	0.15	40.96
32	*ERG3*	2.48	< 0.0001	47.08
33	*ERG3*	1.19	0.11	49.30
34	*ERG3*	1.89	0.08	8.45
35	*ERG5*	2.56	0.33	0.99

File S1

APPEARANCE OF IDENTICAL MUTATIONS

Five identical ergosterol mutations were sampled within multiple lines (Table 1). The most likely explanation is that these mutations arose during population expansion before the lineages were isolated from one another and before the stressor was applied. The culture used to seed mutation acquisition screens was derived from a single wild type colony grown up overnight in YPD, an unstressful environment. Given the size of this overnight culture ($\sim 1.2 \times 10^9$ cells), there would have been approximately 30 generations of growth during this YPD phase ($2^{30} = 1.1 \times 10^9$). Despite the bottleneck to a single colony-forming unit, the population size of the source population, a plausible per-base pair mutation rate (0.33×10^{-9}, LYNCH *et al.* 2008), the hundreds of one-step mutations that could potentially result in nystatin tolerance ($\sim$ 350 different mutations based on our data of what types of mutations confer tolerance to nystatin, see section below), and the number of founding lineages (60 and 180 in screens 'a' and 'b', respectively), we calculated that there is a high probability that some beneficial mutations were segregating in the precursor population (see File S2). Standing genetic variation from a colony grown for a single overnight in YPD has previously been found to play a large role in the eventual mutations that were selected in a yeast experimental evolution project (GRESHAM *et al.* 2008), and we believe this is also the most likely explanation here.

Well-to-well contamination is also possible, yet unlikely. We kept track of where mutations were isolated within the 96 well plates; in no case was the same mutation isolated in neighbouring wells and in multiple cases the same mutation was isolated from different plates within the same screen. In at least the case of BMN11-15, well-to-well contamination is even less likely, as growth was seen in all wells before the first culture was isolated, thus there was little to no opportunity for contamination from one well to another (Table S1).

Although certainly possible biologically, we do not have any support for the same allelic variant arising independently in multiple lines. Mutation lines were acquired in two separate screens (denoted 'a' and 'b' in Table S1). The exact same protocol was used for both acquisition screens (see Methods), however, a different colony was grown up overnight to provide the culture used to seed all replicate wells in each screen. In no case was the same allele sampled in the two different screens.

EFFECT OF NON-ERGOSTEROL MUTATIONS

While we attempted to minimize the number of mutations carried by the lines, secondary mutations could have arisen and fixed during the $\sim$30 generations of growth in YPD within the precursor population or during the $\sim$30 generations of growth in nystatin required for yeast precipitate to be detected. Nine lines carry non-synonymous nuclear mutations in genes not involved in the ergosterol pathway (Table S2). Two of these lines share the same ergosterol mutation

with other BMN lines, which allows us to directly assess the phenotypic affect of these additional mutations. BMN27 carries three additional mutations: nonsynonymous mutations in YJR107W (an uncharacterized protein), *AUR1* (a protein required for sphingolipid synthesis), and an extra copy of chromosome 2. BMN27 has a higher IC_{50} in both salt and copper than the three other lines that carry the same ergosterol mutation (BMN24-26), but these differences are not significant. BMN15, with a nonsynonymous mutation in *MBP1*, also does not differ in our fitness assays from BMN11-14, lines with which it shares *ERG6* and *GDA1* mutations. The remaining lines with secondary mutations have very similar nystatin tolerance to other lines that carry mutations in the same ergosterol gene. As a further test, we backcrossed and sporulated representative lines that contain a mutation in each of the four ergosterol genes (BMN1, BMN9, BMN32 and BMN58) to BY4739. For each backcross we found 2:2 segregation of nystatin tolerance. We thus have little reason to suspect that mutations in non-ergosterol genes are strongly influencing our results.

STATISTICAL RESULTS REMAIN THE SAME IF WE COMBINE LINES WITH THE SAME ERGOSTEROL MUTATION

Statistical results reported in the main text are upheld if we use the average tolerance and fitness results from lines that contain the same ergosterol mutation. For all three assays conducted in nystatin (IC_{50}, growth rate, $OD48$), and both assays conducted in YPD (growth rate and $OD48$), we recover the same results previously reported, only the ergosterol gene that bears a mutation has a significant effect on the results (i.e., mutation type and their interaction do not).

The statistical results of our correlation tests between different fitness proxies also yield the same results if we average across line replicates with the same mutation. Comparing between nystatin tolerance breadth (IC_{50}) and two fitness proxies in nystatin we find that all three assays are significantly correlated to each other (growth rate and $OD48$: cor = 0.88, t_{18} = 7.9, $p < 0.0001$; growth rate and IC_{50}: cor = 0.80, t_{18} = 5.7, $p < 0.0001$; $OD48$ and IC_{50}: cor = 0.72, t_{18}= 4.4,p = 0.0003). When we compare IC_{50} in nystatin and the same two fitness proxies when the lines are grown in YPD we find the same result we previously reported, i.e., the only significant correlation is growth rates in YPD with IC_{50} in nystatin (growth rate and $OD48$ in YPD: cor = 0.35, t_{18} = 1.6, p = 0.13; growth rate in YPD and IC_{50} in nystatin: cor = 0.65, t_{18} = 3.6, p = 0.002; $OD48$ and IC_{50} in nystatin: cor = 0.20, t_{18} = 0.9, p = 0.4).

Finally, we also recover the same pattern of tradeoffs between tolerance to nystatin and secondary environments (all measured as IC_{50}). Specifically, we find tolerance to nystatin is significantly correlated to both ethanol (cor = -0.62, t_{19} = -3.5, p = 0.002) and copper (cor = -0.88, t_{19} = -8.0, $p < 0.0001$), but not to salt (cor = 0.13, t_{19} = 0.6, p = 0.56).

LITERATURE CITED

GRESHAM, D., M. DESAI, C. TUCKER, H. JENQ, D. PAI, *et al.*, 2008 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet **4**: e1000303.

LYNCH, M., W. SUNG, K. MORRIS, N. COFFEY, and C. LANDRY, 2008 A genome-wide view of the spectrum of spontaneous mutations in yeast. PNAS **105**: 9272–9277.

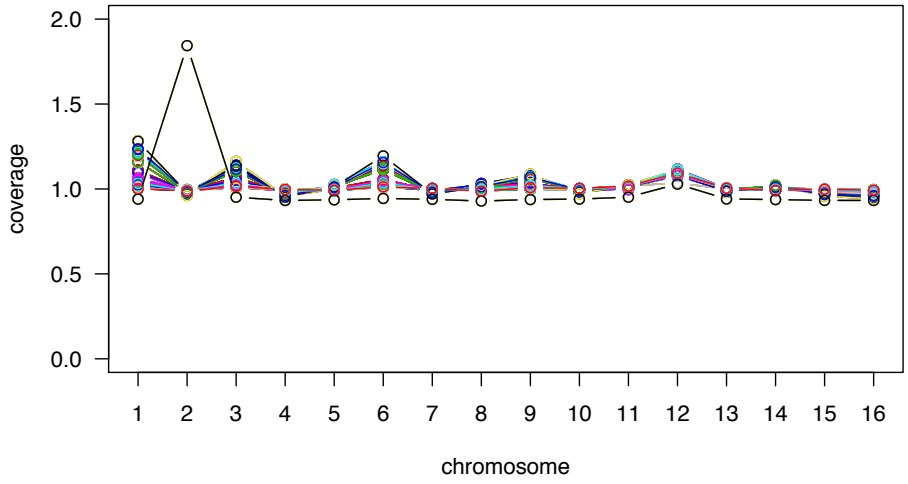


Figure S1 Relative coverage of each chromosome from genomic alignments. Using the Illumina genomic sequence data, the total coverage for each chromosome was calculated as the proportion of sequenced sites mapping to a particular chromosome relative to the proportion of known mapped sites located on that chromosome within the yeast reference genome (as reported by configureBuild.pl in Illumina's CASAVA-1.8.0 package). Examining the coverage data for each chromosome from each BMN line (each line is plotted with a unique colour) indicates only one aneuploidy event - an extra copy of chromosome 2 in BMN27.

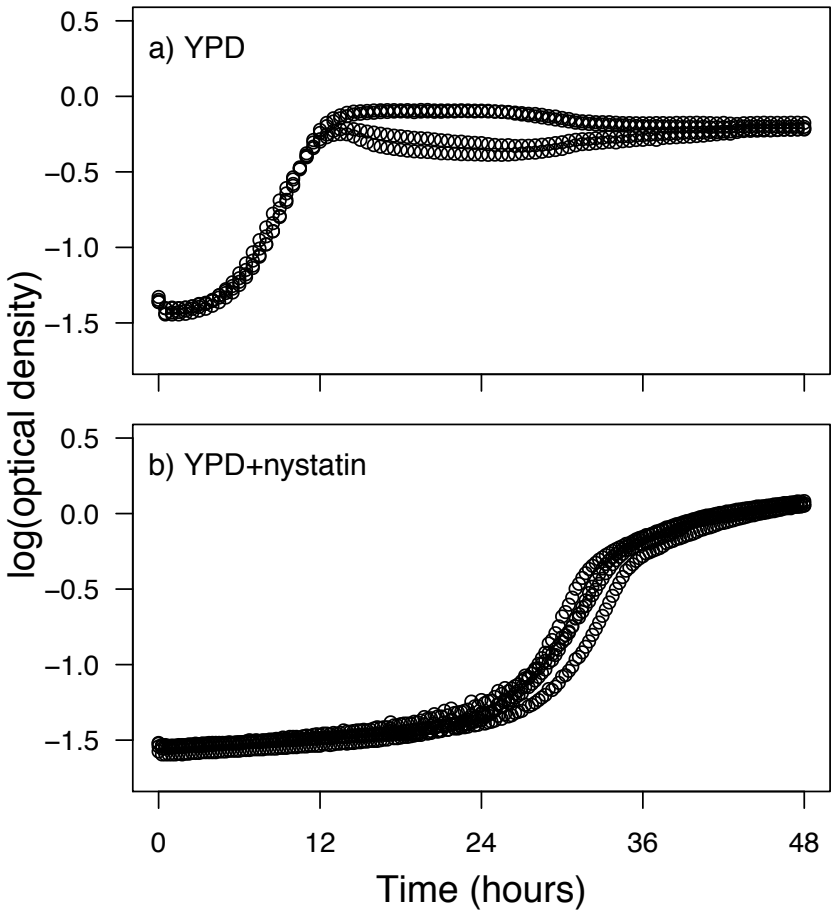


Figure S2 Representative raw growth curves for one line (BMN31) in a) YPD and b) YPD+4μM nystatin. Optical density was measured automatically by a Bioscreen C Microbiology Workstation. Optical density has stopped increasing by 48 hours in YPD, while lines are still growing at 48 hours in nystatin. The optical density at 48 hours ($OD48$) thus reflects primarily resource efficiency in YPD while it represents a combined measure of resource efficiency and growth rate in nystatin.

Table S1 The date mutations were acquired. Mutations were acquired in two screens ('a' & 'b'), with each acquisition experiment lasting one week. Different ancestral colonies were used to initiate each acquisition screen. By examining the date of isolation and screen we gain insight into the process that led to identical ergosterol mutations in multiple lines (BMN2-4, BMN7-10, BMN11-15, BMN17-20 and BMN24-27).

BMN	Date Isolated	Screen
1	10.07.25	a
2	10.08.14	b
3	10.08.14	b
4	10.08.09	b
5	10.07.23	a
6	10.07.25	a
7	10.07.25	a
8	10.07.25	a
9	10.07.25	a
10	10.07.23	a
11	10.08.09	b
12	10.08.09	b
13	10.08.09	b
14	10.08.09	b
15	10.08.09	a
16	10.07.23	a
17	10.07.25	a
18	10.07.25	a
19	10.07.23	a
20	10.07.23	a
21	10.08.10	b
22	10.08.12	b
23	10.08.12	b
24	10.08.10	b
25	10.08.12	b
26	10.08.12	b
27	10.08.13	b
28	10.07.25	a
29	10.07.25	a
30	10.08.10	b
31	10.08.10	b
32	10.08.12	b
33	10.08.10	b
34	10.08.13	b
35	10.08.14	b

Table S2 Genotypic basis of mutations in genes not in the ergosterol biosynthesis pathway. Nineteen additional unique mutations plus one aneuploidy were found within 19 lines.

BMN Line	Gene	Genome Position (Chr.Bp)	Gene Position (in nucleotides)	Mutation	Amino Acid Change
1	YOL073C	XV.193885	916	G>A	Asp306Asn
3	YPL039W	XVI.479629	405	G>A	synonymous (Arg)
5	*CDC23*	VIII.438829	222	A>C	synonymous (Ile)
8		X.30640		A>C	
11-15	*GDA1*	V.74568	798	C>T	synonymous (Val)
15	*MBP1*	IV.354071	1195	T>G	Phe399Val
16		XIV.1753512		A>T	
16		XIV.1753521		C>A	
22	*COX1*	mt.23360	9543	T>A	synonymous (Ile)
22	*COX1*	mt.23361	9544	A>T	Ile3182Leu
23	*SCW11*	VII.442319	591	A>G	synonymous (Ser)
24-27	*FCY2*	V.267873	241	G>A	Glu81Lys
26		XIV.507563		T>G	
27	*AUR1*	XI.436609	1030	C>T	Pro344Ser
27	YJR107W	X.627995	656	G>A	Trp219Stop
27				+Chr2	
29	*SGS1*	XIII.644130	1129	A>G	Asn377Asp
30	*MDM20*	XV.188973	1950	T>G	Ile650Met
31	*ALT1*	XII.319765	251	T>C	Leu84Pro
35		I.73925		A>G	

Table S3 Maximum likelihood results for growth in nystatin. We fit a likelihood model to the combined data for each mutation line and the ancestral line. The full model allowed two IC_{50} values to be fit to the data, while the constrained model forced the IC_{50} for the mutant and ancestral lines to be equal. All mutation lines were found to have significantly different IC_{50} values from the ancestor (difference in log-likelihood between the two models > 1.92; see Methods).

BMN	Gene	LogLikelihood, FullModel	LogLikelihood, Constrained	Difference in Models
	*erg6*Δ	35.60	-12.04	47.64
	*erg5*Δ	18.18	-10.64	28.82
	*erg3*Δ	46.23	5.76	40.46
1	*ERG7*	18.12	10.77	7.34
2	*ERG6*	50.14	-5.47	55.61
3	*ERG6*	38.97	-16.94	55.91
4	*ERG6*	60.91	-18.95	79.86
5	*ERG6*	37.31	-13.68	50.99
6	*ERG6*	60.87	-0.47	61.34
7	*ERG6*	53.25	-2.56	55.81
8	*ERG6*	49.38	-16.30	65.68
9	*ERG6*	35.15	-16.92	52.08
10	*ERG6*	42.37	-13.91	56.28
11	*ERG6*	62.46	6.72	55.74
12	*ERG6*	63.21	7.17	56.04
13	*ERG6*	69.14	12.87	56.27
14	*ERG6*	38.78	-16.81	55.59
15	*ERG6*	51.53	-17.35	68.89
16	*ERG6*	39.39	-15.73	55.12
17	*ERG6*	68.80	12.80	56.00
18	*ERG6*	35.34	-13.35	48.69
19	*ERG6*	35.71	-16.51	52.22
20	*ERG6*	42.61	-12.49	55.10
21	*ERG3*	43.94	8.56	35.37
22	*ERG3*	33.01	-0.17	33.18
23	*ERG3*	31.41	-0.68	32.09
24	*ERG3*	31.01	6.83	24.18
25	*ERG3*	28.70	2.33	26.37
26	*ERG3*	57.92	14.94	42.98
27	*ERG3*	60.53	23.06	37.47
28	*ERG3*	34.21	-3.92	38.13
29	*ERG3*	41.40	4.49	36.91
30	*ERG3*	48.29	17.20	31.09
31	*ERG3*	29.86	-0.50	30.36
32	*ERG3*	35.36	9.62	25.74
33	*ERG3*	29.21	-0.11	29.32
34	*ERG3*	25.92	1.46	24.46
35	*ERG5*	-3.30	-14.00	10.70

Table S4 T-test results comparing growth rate of BMN lines in nystatin to five ancestral colonies.

BMN	t	df	p-value
1	-3.4	18.7	0.003
2	-15.5	35.8	< 0.0001
3	-4.1	6.9	0.005
4	-4.1	6.8	0.005
5	-7.1	8.7	< 0.0001
6	-16	33.6	< 0.0001
7	-5.9	9.3	0.0002
8	-12.4	20.3	< 0.0001
9	-7.6	14.6	< 0.0001
10	-4.1	6.1	0.006
11	-16.8	33.6	< 0.0001
12	-7.6	8.6	< 0.0001
13	-9.1	7.1	< 0.0001
14	-15.8	30.8	< 0.0001
15	-6.6	7.5	0.0002
16	-2.4	3.3	0.087
17	-5.4	8.4	0.0005
18	-4.0	9.2	0.003
19	-6.7	7.6	0.0002
20	-17	32.5	< 0.0001
21	-1.7	3.4	0.18
22	-3.1	5.3	0.026
23	-2.4	5.2	0.06
24	-4.4	7.8	0.002
25	-3.1	10.3	0.011
26	-6.2	23.0	< 0.0001
27	-2.6	6.8	0.034
28	-3.5	3.7	0.027
29	-2.5	6.0	0.046
30	-4.1	6.5	0.0056
31	-2.9	5.5	0.029
32	-2.9	7.2	0.021
33	-5.1	33.8	< 0.0001
34	-5.4	7.2	0.0009
35	-2.1	4.4	0.10

Table S5 T-test results comparing $OD48$ of BMN lines grown in nystatin to five ancestral colonies.

BMN	t	df	*p*-value
1	-5.5	8.2	0.0005
2	-10.8	5.2	0.0001
3	-8.1	5.1	0.0004
4	-7.1	5.1	0.0008
5	-10.5	5.1	0.0001
6	-12.1	3.1	0.001
7	-9.2	6.2	0.0001
8	-13.7	3.1	0.0007
9	-8.7	8.2	< 0.0001
10	-4.5	5.0	0.0066
11	-9.8	3.1	0.002
12	-10.9	5.2	0.0001
13	-14.0	3.2	0.0006
14	-6.4	2.0	0.023
15	-8.4	5.1	0.0004
16	-2.6	3.0	0.082
17	-5.9	6.1	0.001
18	-5.3	7.1	0.001
19	-6.1	5.1	0.002
20	-11.3	4.1	0.0003
21	-2.8	3.0	0.065
22	-4.4	4.1	0.012
23	-3.8	4.0	0.018
24	-7.2	5.2	0.0007
25	-8.2	6.2	0.0001
26	-8.8	4.3	0.0007
27	-4.4	5.1	0.007
28	-3.8	3.0	0.032
29	-3.0	5.1	0.030
30	-6.0	5.1	0.002
31	-10.8	4.4	0.0003
32	-7.0	5.2	0.0007
33	-6.1	3.1	0.008
34	-8.8	5.1	0.0003
35	-2.3	4.0	0.087

Table S6 T-test results comparing growth rate of BMN lines grown in YPD to five ancestral colonies.

BMN	t	df	*p*-value
1	6.0	20.6	< 0.0001
2	4	20.1	0.0008
3	3.7	13.7	0.0026
4	4.7	21.0	0.0001
5	5	20.4	0.0001
6	3.3	21.6	0.0033
7	3.5	21.8	0.0019
8	3.2	7.0	0.0146
9	3.9	21.8	0.0007
10	5.8	21.8	< 0.0001
11	3.9	5.2	0.011
12	4.7	18.0	0.0002
13	3.7	9.1	0.0045
14	4.1	19.9	0.0006
15	4.9	20.4	0.0001
16	4.9	10.1	0.0006
17	3.9	17.0	0.0012
18	4.4	22.0	0.0002
19	5.0	20.1	0.0001
20	5.5	19.4	< 0.0001
21	6.8	7.2	0.0002
22	14.3	21.8	< 0.0001
23	14.0	22.0	< 0.0001
24	6.8	9.3	0.0001
25	10.9	12.6	< 0.0001
26	13.4	21.3	< 0.0001
27	5.6	11.1	0.0002
28	10.1	12.9	< 0.0001
29	5.1	8.8	0.0007
30	7.6	14.1	< 0.0001
31	5.1	6.6	0.0017
32	12.6	21.6	< 0.0001
33	4.5	6.7	0.0029
34	7.5	7.4	0.0001
35	2.6	11.4	0.022

Table S7 T-test results comparing $OD48$ of BMN lines grown in YPD to five ancestral colonies.

BMN	t	df	*p*-value
1	5.5	12.0	0.0001
2	6.0	19.0	<0.0001
3	3.8	4.0	0.020
4	12.2	7.9	< 0.0001
5	3.5	4.0	0.024
6	4.5	4.0	0.011
7	10.1	5.9	0.0001
8	7.3	7.7	0.0001
9	12.3	9.5	< 0.0001
10	5.3	4.9	0.003
11	5.2	5.9	0.0022
12	4.7	4.9	0.0058
13	7.0	8.2	0.0001
14	3.1	3.4	0.043
15	8.8	11.5	< 0.0001
16	13.7	21.8	< 0.0001
17	3.6	3.7	0.026
18	2.8	3.3	0.063
19	4.4	3.4	0.016
20	3.5	3.8	0.027
21	10.1	5.7	0.0001
22	17.8	21.7	< 0.0001
23	4.3	3.7	0.015
24	6.2	5.9	0.0009
25	8.0	5.5	0.0003
26	9.2	8.6	< 0.0001
27	10.0	6.0	0.0001
28	10.6	9.5	< 0.0001
29	7.8	3.9	0.0016
30	7.7	5.0	0.0006
31	18.4	20.4	< 0.0001
32	5.0	4.5	0.0052
33	15.7	22.0	< 0.0001
34	5.7	7.6	0.0006
35	2.3	4.1	0.079

Table S8 Likelihood ratio tests comparing IC_{50} of ancestral and BMN lines in copper, ethanol and salt.

		Copper (CuSO4)	Ethanol	Salt (NaCl)
BMN	Gene	Difference in Models	Difference in Models	Difference in Models
1	*ERG7*	3.76	30.00	4.79
2	*ERG6*	5.30	11.75	23.87
3	*ERG6*	0.62	5.18	23.50
4	*ERG6*	2.15	12.06	27.00
5	*ERG6*	1.73	24.35	26.36
6	*ERG6*	12.24	25.31	19.81
7	*ERG6*	1.93	14.42	20.05
8	*ERG6*	2.58	17.05	26.35
9	*ERG6*	2.61	17.76	30.26
10	*ERG6*	3.05	15.25	34.54
11	*ERG6*	2.99	8.05	18.08
12	*ERG6*	2.39	4.19	25.83
13	*ERG6*	1.88	18.78	32.90
14	*ERG6*	13.63	17.55	27.48
15	*ERG6*	1.91	29.10	29.00
16	*ERG6*	4.34	20.17	26.91
17	*ERG6*	0.44	6.95	23.83
18	*ERG6*	3.49	21.75	27.88
19	*ERG6*	1.21	20.40	24.13
20	*ERG6*	1.94	33.79	28.62
21	*ERG3*	2.20	0.47	48.55
22	*ERG3*	2.00	0.02	49.20
23	*ERG3*	2.50	1.00	48.42
24	*ERG3*	10.06	0.02	9.90
25	*ERG3*	2.17	0.15	40.52
26	*ERG3*	5.49	0.48	40.25
27	*ERG3*	9.99	0.49	0.44
28	*ERG3*	5.48	0.58	45.23
29	*ERG3*	9.82	< 0.0001	45.53
30	*ERG3*	1.18	0.01	44.65
31	*ERG3*	3.20	0.15	40.96
32	*ERG3*	2.48	< 0.0001	47.08
33	*ERG3*	1.19	0.11	49.30
34	*ERG3*	1.89	0.08	8.45
35	*ERG5*	2.56	0.33	0.99

Here we model the growth of a population from a single cell established on an agar plate (YPD), picked as a colony, and grown to saturation in 10mL YPD, which corresponds to a population size of ~10^9cells ("source" population), from which a sample of cells is taken to establish each individual lineage (the "founding" population).

Parameters:

μ = total mutation rate to nystatin resistance (0.33 ◊ 10^{-9}(Lynch et al., 2008, PNAS); range explored: $10^{-9} - 10^{-6}$)

N1 = population size at saturation in 10mL YPD (1.2×10^9 cells, measured)

f = fraction of population sampled to found a lineage (0.001 = 10 ul/10 mL)

L = number of founding lineages started from the YPD culture (note there are two mutation accumulation "screens", the first with 60 well replicates, the second with 180)

lowerMUT = lower estimate of the target size

upperMUT = upper estimate of the target size

```
tryμ = 0.33 × 10^-9;
tryN1 = 1.2 * 10^9;
tryf = 0.001;
tryL1 = 60;
tryL2 = 180;
lowerMUT = 246;
upperMUT = 522;
```

Calculation #1: Estimating the genome-wide target size for mutations to nystatin resistance in our screens

Here, we estimate the number of sites that could potentially yield nystatin resistance by a single basepair mutation. Because multiple resistant erg3 and erg6 mutations were obtained and many of these created a stop codon, we first ask how many nucleotides in these two genes could mutate to a stop codon. We then use the observed proportion of unique mutations that were stop codons in erg3 and erg6 to correct this etimate and obtain an overall estimate of the number of sites that could potentially yield nystatin resistance. While the estimation is rough, it gives a reasonable expectation that there are hundreds of potential sites in the genome that could give rise to nystatin resistance of the nature observed in our experiment (many additional minor effect mutations are possible but were not obtained by our assay. Note that then consider how this number might be an over- or under-estimate.

Copy and paste sequences of ERG3 and ERG6 (from yeastgenome.org)

```
erg3 =
  "ATGGATTTGGTCTTAGAAGTCGCTGACCATTATGTCTTAGACGACTTGTACGCTAAAGTTCTGCCCGCTTCGTTGGCAGCTAATA\
     TTCCTGTCAAGTGGCAGAAATTGCTAGGGTTGAACAGTGGGTTCAGCAATTCTACGATTTTGCAGGAGACTTTGAACTCC\
     AAGAATGCCGTCAAAGAATGTAGAAGGTTCTACGGGCAGGTGCCATTCCTGTTTGATATGTCGACGACGTCTTTTGCATC\
     GCTATTGCCTCGTTCCAGCATCTTGAGAGAATTCCTCTCACTATGGGTTATTGTTACGATCTTTGGTTTACTACTTTACT\
     TATTCACGGCTAGTCTCAGCTACGTGTTTGTGTTTGACAAGTCGATTTTCAACCATCCTCGTTACTTGAAAAACCAAATG\
     GCAATGGAAATCAAGTTGGCAGTCAGTGCTATCCCATGGATGTCGATGTTGACCGTTCCATGGTTTGTTATGGAATTGAA\
     CGGCCATTCTAAACTATACATGAAGATTGATTATGAAAACCACGGTGTAAGGAAGCTCATTATCGAGTACTTCACTTTCA\
     TCTTTTTCACTGATTGCGGTGTGTATTTAGCGCACAGATGGTTGCATTGGCCAAGGGTCTACCGTGCTCTGCACAAGCCT\
     CATCACAAGTGGCTGGTCTGCACACCTTTCGCATCTCATTCTTTCCATCCTGTAGACGGGTTTTTGCAATCCATCTCGTA\
     CCACATCTACCCATTGATTCTGCCATTACACAAGGTTTCTTATTTGATTCTGTTCACTTTTGTTAACTTTTGGACTGTTA\
     TGATTCATGACGGTCAATACCTATCAAACAATCCTGCCGTCAACGGTACTGCCTGCCACACGGTTCACCATCTATATTTC\
     AACTACAACTACGGTCAATTCACCACTCTGTGGGACAGACTAGGGGGTTCTTACCGTAGACCAGATGACTCATTGTTTGA\
     TCCTAAGTTAAGAGATGCTAAGGAGACCTGGGACGCTCAAGTTAAGGAAGTTGAACATTTCATCAAGGAGGTCGAAGGTG\
     ATGATAATGATAGAATCTATGAAAACGACCCAAATACCAAGAAGAACAACTGA";

erg6 =
  "ATGAGTGAAACAGAATTGAGAAAAAGACAGGCCCAATTCACTAGGGAGTTACATGGTGATGATATTGGTAAAAAGACAGGTTTGA\
     GTGCATTGATGTCGAAGAACAACTCTGCCCAAAAGGAAGCCGTTCAGAAGTACTTGAGAAATTGGGATGGTAGAACCGAT\
     AAAGATGCCGAAGAACGTCGTCTTGAGGATTATAATGAAGCCACACATTCCTACTATAACGTCGTTACAGATTTCTATGA\
     ATATGGTTGGGGTTCCTCTTTCCATTTCAGCAGATTTTATAAAGGTGAGAGTTTCGCTGCCTCGATAGCAAGACATGAAC\
     ATTATTTAGCTTACAAGGCTGGTATTCAAAGAGGCGATTTAGTTCTCGACGTTGGTTGTGGTGTTGGGGGCCCAGCAAGA\
     GAGATTGCAAGATTTACCGGTTGTAACGTCATCGGTCTAAACAATAACGATTACCAAATTGCCAAGGCAAAATATTACGC\
     TAAAAAATACAATTTGAGTGACCAAATGGACTTTGTAAAGGGTGATTTCATGAAAATGGATTTCGAAGAAAACACTTTCG\
     ACAAAGTTTATGCAATTGAGGCCACATGTCACGCTCCAAAATTAGAAGGTGTATACAGCGAAATCTACAAGGTTTTGAAA\
     CCGGGTGGTACCTTTGCTGTTTACGAATGGGTAATGACTGATAAATATGACGAAAACAATCCTGAACATAGAAAGATCGC\
     TTATGAAATTGAACTAGGTGATGGTATCCCAAAGATGTTCCATGTCGACGTGGCTAGGAAAGCATTGAAGAACTGTGGTT\
     TCGAAGTCCTCGTTAGCGAAGACCTGGCGGACAATGATGATGAAATCCCTTGGTATTACCCATTAACTGGTGAGTGGAAG\
     TACGTTCAAAACTTAGCTAATTTGGCCACATTTTTCAGAACTTCTTACTTGGGTAGACAATTTACTACAGCAATGGTTAC\
     TGTAATGGAAAAATTAGGTCTAGCCCCAGAAGGTTCCAAGGAAGTTACTGCTGCTCTAGAAAATGCTGCGGTTGGTTTAG\
     TTGCCGGTGGTAAGTCCAAGTTATTCACTCCAATGATGCTTTTCGTCGCTAGGAAGCCAGAAAACGCCGAAACCCCCTCC\
     CAAACTTCCCAAGAAGCAACTCAATAA";
```

Calculate the number of codons
there are 366 codons in erg3
there are 384 codons in erg6

```
lenERG3 = Length[Characters[erg3]] / 3
```

```
366
```

```
lenERG6 = Length[Characters[erg6]] / 3
```

```
384
```

Split the strings into codons

```
codonsERG3 =
 Table[StringJoin[Table[Characters[erg3][[j]], {j, 3 * i - 2, 3 * i}]], {i, 1, lenERG3}]
codonsERG6 =
 Table[StringJoin[Table[Characters[erg6][[j]], {j, 3 * i - 2, 3 * i}]], {i, 1, lenERG6}]
```

The stop codons are: UAA, UAG, UGA (TAA, TAG, TGA)

Thus, the following 27 codons can mutate to a stop codon by a single basepair change:

```
{"CAA", "GAA", "AAA", "TGA", "TCA", "TTA", "TAG", "TAC", "TAT",
  "CAG", "GAG", "AAG", "TGG", "TCG", "TTG", "TAA", "TAC", "TAT",
  "CGA", "GGA", "AGA", "TAA", "TCA", "TTA", "TGG", "TGC", "TGT"};
```

We can exclude 4 cases involving stop codons:

```
{"CAA", "GAA", "AAA", "TCA", "TTA", "TAC", "TAT",
  "CAG", "GAG", "AAG", "TGG", "TCG", "TTG", "TAC", "TAT",
  "CGA", "GGA", "AGA", "TCA", "TTA", "TGG", "TGC", "TGT"};
```

Among these, the following codons can be hit in two places to yield a stop codon:

```
twice = {"TTA", "TCA", "TAC", "TAT", "TGG"};
```

The following codons can be hit in one (and only one) place to yield a stop codon:

```
once = {"CAA", "GAA", "AAA", "CAG", "GAG",
    "AAG", "TCG", "TTG", "CGA", "GGA", "AGA", "TGC", "TGT"};
```

Thus, the total number of sites that can mutate in erg3 or erg6 to a stop codon is:
#157 basepairs could mutate to create a stop codon in one step from erg3
#176 basepairs could mutate to create a stop codon in one step from erg3

```
Total[StringCount[codonsERG3, once]] +
 2 * Total[StringCount[codonsERG3, twice]]
```

```
157
```

```
Total[StringCount[codonsERG6, once]] +
 2 * Total[StringCount[codonsERG6, twice]]
```

```
176
```

This gives a total number of sites in erg3 and erg6 where a single mutation could generate a stop codon:

```
total = % + %%
```

```
333
```

Some stop codons, however, might not yield a nystatin-resistant phenotype, i.e., those near the end of the gene.
In our dataset, the last sampled stop codon in erg6 was at amino acid 223 and the last sampled erg3 stop was at amino acid 299.
Limiting our counts of potential stop codon hits to this point yields:

```
Total[StringCount[Take[codonsERG3, 299], once]] +
 2 * Total[StringCount[Take[codonsERG3, 299], twice]]
```

```
126
```

```
Total[StringCount[Take[codonsERG6, 223], once]] +
 2 * Total[StringCount[Take[codonsERG6, 223], twice]]
```

```
109
```

This gives a total number of sites in erg3 and erg6 where a single mutation could generate a stop codon at a position before or at the last non-sense mutation observed:

```
total = % + %%
```

```
235
```

Thus, there are at likely to be at least 235 possible one-step stop codons in erg3 & erg6 that could confer the tolerance phenotype.

In our data set, 9 out of 20 unique mutations we found were stop codons in erg3 or erg69. If we use this fraction also for the unobserved mutations, we obtain an estimate of:

```
235 / (9 / 20) // N
```

```
522.222
```

The above suggests that the genome-wide target size for mutations that would generate sufficient nystatin resistance to allow growth in our assays is somewhere in the range of 246 - 522 (the lower bound set by assuming that the only hits that we missed were other one-step stop codons, plus previous hits: 235+11; the upper bound set by assuming that we would have the same fraction of one-step stop codons in the unobserved mutations as we did in the observed mutations).

■ Calculation #2: Chance that the source population contains no nystatin mutations

Number of cell cycles required to produce source population:

```
cycles = Log[2, N1]
```

$$\frac{\text{Log}[N1]}{\text{Log}[2]}$$

Total number of cell divisions involved (1 cell division from 1 -> 2 cells, 2 cell divisions from 2 -> 4 cells, etc):

```
divisions = Sum[2^i, {i, 0, cycles - 1}]
```

$$-1 + N1$$

For example, to go from 1 -> 4 cells involves a total of 3 dividing cells (-> 8 cells would involve 7 dividing cells: one 1->2, two 2->4, and four 4->8):

```
divisions /. N1 → {4, 8}
```

$$\{3, 7\}$$

```
divisions /. N1 → tryN1 // N
```

$$1.2 \times 10^9$$

The chance that NONE of these cell divisions involved a mutant is:

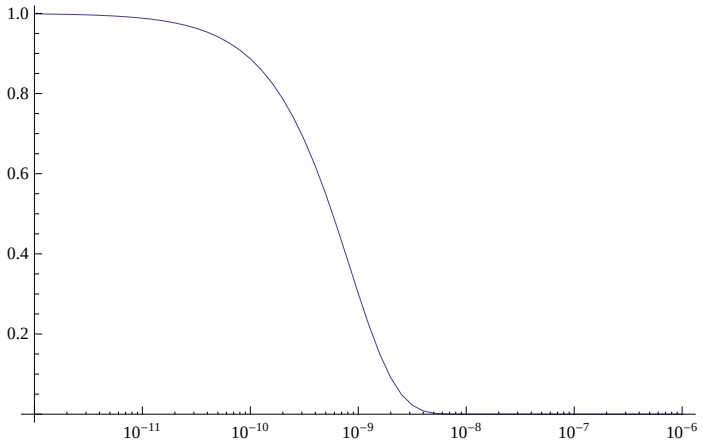
```
(1 - μ)^divisions;
```

Given that μ is small, this is approximately:

```
nomutantsinsource[μ_, N1_] = e^(-μ*N1);
```

Assuming a population size of 1.2×10^9, this probability declines rapidly as the per-basepair mutation rate rises above 10^{-10}:

```
ListLogLinearPlot[
 Table[{10^i, nomutantsinsource[10^i, tryN1]}, {i, -12, -6, 0.1}], Joined → True]
```



Using a per-basepair mutation rate of $0.33 \lozenge 10^{-9}$(Lynch et al., 2008, PNAS), the chance of no mutations at a single site within the source pool would be:

```
nomutantsinsource[tryμ, tryN1]
```

$$0.673007$$

Given potentially hundreds of sites within ERG genes, the chance that none of these sites was polymoprhic in the precursor population is:

```
nomutantsinsource[tryμ × lowerMUT, tryN1]
```

4.92911×10^{-43}

```
nomutantsinsource[tryμ × upperMUT, tryN1]
```

1.68314×10^{-90}

In other words, we're virtually certain that there would be a segregating mutation at some ERG gene within the precursor population.

This does not tell us, however, how likely we are to found multiple lineages carrying the same mutation. For that, we need to determine the frequency of these mutations (many of which will have appeared in the last cell division, thus present in a single cell, and cannot found multiple lineages).

■ Calculation #3 - Calculating the fraction of the source population carrying a mutation

For a mutation rate of $0.33 \lozenge 10^{-9}$, the expected number of mutations at a specific site across all of the cells and cell divisions is approximately:

```
tryμ * divisions /. N1 → tryN1
```

0.396

The number of cell cycles required to go from one cell (original cell that lead to a colony on a plate) to 1.2×10^{9} cells (the number after letting one colony grow overnight in 10mL YPD)

```
Floor[cycles /. N1 → tryN1]
```

30

Thus, we expect less than one hit in total across all of the ~30 cell cycles at any one particular site.

Next, we derive the probability that a single mutation hits in any particular cell cycle.

At one specific site, the probability that a mutation occurs in the kth cell cycle (going from 2^{k-1} cells to 2^{k} cells) is:

```
prob[k_, μ_, N1_] = 1 - (1 - μ)^n /. n → 2^(k-1);
```

based on one minus the probability that no mutation hits. (Technically, this allows for the possibility that more than one hit would occur, but as we saw above, there is unlikely to be more than one mutation across all cell divisions and so we don't expect more than one mutation in any one cell cycle. In the following, we assume that if there is a hit in a particular cell cycle, it creates only a single daughter mutant cell.)

If a mutation does occur in the kth cycle (i.e., among the 2^{k} cells that result in this cycle, one is a new mutant), the fraction of the final population that will be mutant is:

```
frac[k_, μ_, N1_] = 1 / 2^k;
```

This gives us the probability distribution for the fraction of mutant cells in the source population (amounting to a number of mutant cells: $n = N1 / 2^{k}$), where the probability of no mutant cells at a particular site equalling:

```
1 - Sum[prob[k, μ, N1], {k, 1, cycles}]
```

$$1 - \sum_{k=1}^{\frac{\text{Log}[N1]}{\text{Log}[2]}} \left(1 - (1 - \mu)^{2^{-1+k}}\right)$$

Given L lineages started from the source population (here we use L = 80 for trial 1 and 180 for trial 2) where each lineage is started with a fraction, f, of the source population, we next calculate the probability that two or more will sample the same mutant, if that mutant consists of n cells in the source population:

```
Rehit[n_, N1_, L_, f_] = 1 - ((1 - n/N1)^(f N1))^L -
  Sum[Product[(1 - n/(N1 - (j - 1) * f * N1))^(f N1), {j, 1, i - 1}] * (1 - (1 - n/(N1 - (i - 1) * f * N1))^(f N1)) *
    Product[(1 - (n - 1)/(N1 - (j - 1) * f * N1))^(f N1), {j, i + 1, L}], {i, 1, L}]
```

$$1-\left(\left(1-\frac{n}{N1}\right)^{f\,N1}\right)^{L}-\sum_{i=1}^{L}\left(1-\left(1-\frac{n}{N1-f\,(-1+i)\,N1}\right)^{f\,N1}\right)\left(\frac{\text{Pochhammer}\left[\frac{n-N1}{f\,N1},\,-1+i\right]}{\text{Pochhammer}\left[-\frac{1}{f},\,-1+i\right]}\right)^{f\,N1}\left(\frac{\text{Pochhammer}\left[\frac{-1+n+(-1+f\,i)\,N1}{f\,N1},\,-i+L\right]}{\text{Pochhammer}\left[-\frac{1}{f}+i,\,-i+L\right]}\right)^{f\,N1}$$

This is calculated as one minus the probability that the none of the L lines get the mutant minus the probability that exactly one line gets a mutant. The latter accounts for the fact that each time a lineage is sampled, there are fewer cells in the source population and that any of the L lineages could be the one that gets the hit (sampling without replacement).

Summing over the probability distribution of when the mutation could arise we have:

```
ChanceRehit[μ_, N1_, L_, f_, m_] = Sum[prob[k, μ, N1] * Rehit[N1 / 2^k, N1, L, f], {k, 1, m}]
```

$$\sum_{k=1}^{m}\left(1-(1-\mu)^{2^{-1+k}}\right)\left(1-\left(\left(1-2^{-k}\right)^{f\,N1}\right)^{L}-\sum_{i=1}^{L}\left(1-\left(1-\frac{2^{-k}\,N1}{N1-f\,(-1+i)\,N1}\right)^{f\,N1}\right)\left(\frac{\text{Pochhammer}\left[\frac{-N1+2^{-k}\,N1}{f\,N1},\,-1+i\right]}{\text{Pochhammer}\left[-\frac{1}{f},\,-1+i\right]}\right)^{f\,N1}\left(\frac{\text{Pochhammer}\left[\frac{-1+2^{-k}\,N1+(-1+f\,i)\,N1}{f\,N1},\,-i+L\right]}{\text{Pochhammer}\left[-\frac{1}{f}+i,\,-i+L\right]}\right)^{f\,N1}\right)$$

- **For Trial 1 (L=60)**

For our parameters, at one specific site, the chance of seeing multiple lineages carrying the mutation in the first screen is:

```
ChanceRehit[tryμ, 2^30, tryL1, tryf, 30]
```

```
0.0128062
```

Given potentially hundreds of such sites, however, the chance that at least one of them them would lead to a multiple hit is very high (>95%):

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL1, tryf, 30])^lowerMUT
```

```
0.958025
```

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL1, tryf, 30])^upperMUT
```

```
0.998803
```

Even if we say that the 20 sites that we observed as nystatin resistance represents the entire target size, there is still a decent chance that we would see multiple hits:

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL1, tryf, 30])^20
```

```
0.227234
```

- **For Trial 2 (L=180 wells)**

For our parameters, at one specific site, the chance of seeing multiple lineages carrying the mutation in the second screen is:

```
ChanceRehit[tryμ, 2^30, tryL2, tryf, 30]
```

```
0.0316935
```

Given potentially hundreds of such sites, however, the chance that at least one of them them would lead to a multiple hit is very high (>99.9%):

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL2, tryf, 30])^lowerMUT
```

```
0.999638
```

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL2, tryf, 30])^upperMUT
```

```
1.
```

Even if we say that the 20 sites that we observed as nystatin resistance represents the entire target size, there is still a decent chance that we would see multiple hits:

```
1 - (1 - ChanceRehit[tryμ, 2^30, tryL2, tryf, 30])^20
```

```
0.474882
```

Calculation #4 - Calculating the chance that all sampled lineages would contain a mutation

We can also constrain our parameters by the observation that not all samples carry the same mutation. Here, we use the above calculations to determine the probability that all L lineages would carry the same mutational hit.

With L lineages (each started with a fraction of the source population, f), the probability that all sample a given mutation, if that mutant consists of n cells in the source population, is:

```
ALLhit[n_, N1_, L_, f_] = Product[(1 - (1 - n/N1)^(f N1)), {j, 1, L}]
```

$$\left(1-\left(1-\frac{n}{N1}\right)^{f\,N1}\right)^{L}$$

(This isn't exact, as we should adjust the fraction of mutant cells remaining due to sampling, but the answer will be close.)

Summing over the probability distribution of when the mutation could arise we have:

```
ChanceALLhit[μ_, N1_, L_, f_, m_] = Sum[prob[k, μ, N1] * ALLhit[N1 / 2^k, N1, L, f], {k, 1, m}]
```

$$\sum_{k=1}^{m}\left(1-\left(1-2^{-k}\right)^{f\,N1}\right)^{L}\left(1-(1-\mu)^{2^{-1+k}}\right)$$

- **For Trial 1 (L=60 wells)**

For our parameters, at one specific site, the chance of seeing all L=60 lineages carry a specific mutation in the first screen is:

```
ChanceALLhit[tryμ, 2^30, tryL1, tryf, 30]
```

```
0.0000587235
```

Even with hundreds of such sites, the chances that all lineages would be hit by the same mutation is slim (<5%):

```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL1, tryf, 30])^lowerMUT
```

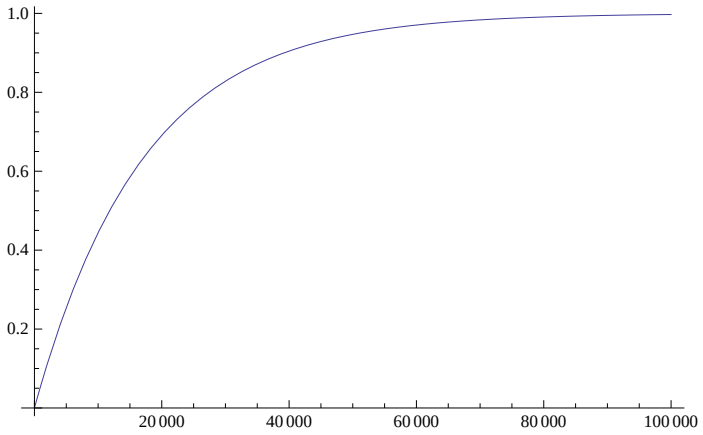
```
0.0143426
```

```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL1, tryf, 30])^upperMUT
```

```
0.0301895
```

We would, however, almost certainly sample the same hit in all L=60 lineages if the target size were >~50,000 sites:

```
Plot[1 - (1 - ChanceALLhit[tryμ, 2^30, tryL1, tryf, 30])^n, {n, 1, 100 000}]
```




```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL1, tryf, 30])^50 000
```

```
0.946936
```

Given that we did not see all of the lineages with the same hit, the above tells us that the target size could not be this large.

- **For Trial 2 (L=180 wells)**

Similarly, the chance of seeing all L=180 lineages carry a specific mutation is:

```
ChanceALLhit[tryμ, 2^30, tryL2, tryf, 30]
```

```
0.0000443119
```

Even with hundreds of such sites, the chances that all lineages would be hit by the same mutation is slim (<5%):

```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL2, tryf, 30])^lowerMUT
```

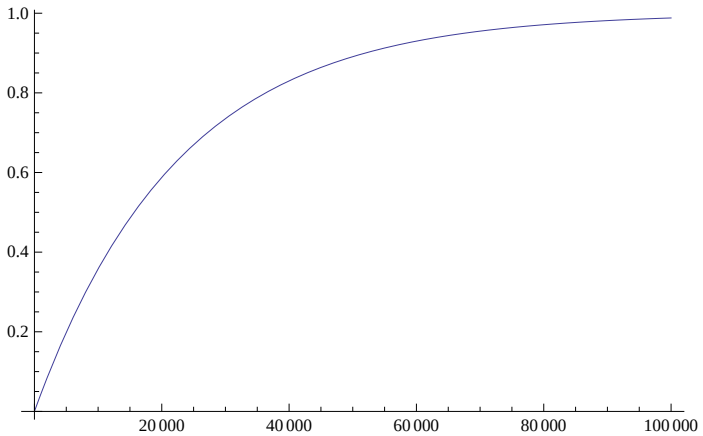
```
0.0108418
```

```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL2, tryf, 30])^upperMUT
```

```
0.0228658
```

We would, however, almost certainly sample the same hit in all L=180 lineages if the target size were >~50,000 sites:

```
Plot[1 - (1 - ChanceALLhit[tryμ, 2^30, tryL2, tryf, 30])^n, {n, 1, 100 000}]
```



```
1 - (1 - ChanceALLhit[tryμ, 2^30, tryL2, tryf, 30])^50 000
```

```
0.890917
```

Given that we did not see all of the lineages with the same hit, the above tells us that the target size could not be this large.

Conclusion

The above calculations inform us that there are likely to be hundreds of potential target sites leading to nystatin resistance (not tens of thousands) and that the chance that the same mutation, occurring at any one of these sites during the clonal expansion of the precursor population, would be sampled in more than one lineage is high (>95% in each screen).