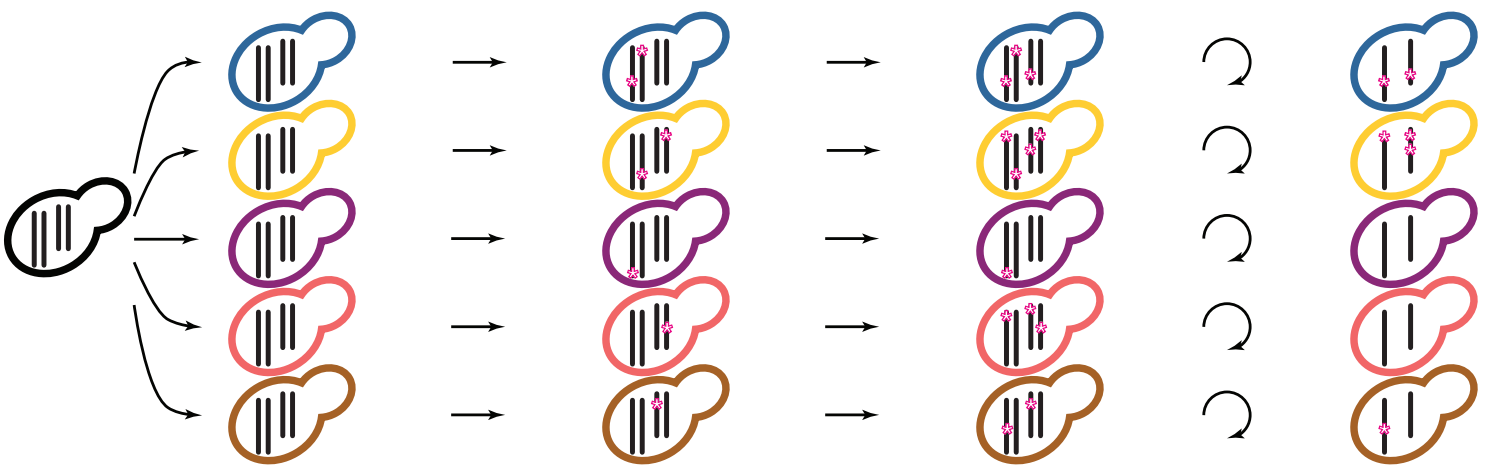
Inferring the properties of mutational effects on fitness using high-throughput phenotyping

Y. Plavskin[1,2], M.S. de Biase[3], N. Ziv[1,2], Y.O. Zhu[4], B. Goulet[1,2], D.W. Hall[5], R.F. Schwarz[3], D.A. Petrov[4], D. Tranchina[2,6], M.L. Siegal[1,2]

1 Center for Genomics and Systems Biology, New York University, New York, NY, USA
2 Department of Biology, New York University, New York, NY, USA
3 Berlin Institute for Medical Systems Biology, Max Delbruck Center for Molecular Medicine, Berlin, Germany;
4 Department of Biology, Stanford University, Stanford, CA, USA
5 Department of Genetics, University of Georgia, Athens, GA, USA
6 Courant Institute for Mathematical Sciences, New York University, New York, NY, USA

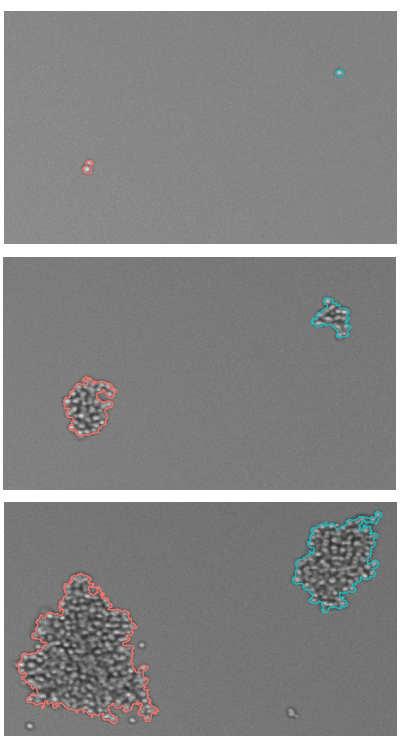
Spontaneous mutations are the source of novel genetic variation in populations, providing the raw material on which selection acts. To shed light on the spectrum of mutations that constitute the starting point of selection and their typical effects on fitness, we estimated the distribution of effect sizes of spontaneous mutations affecting growth in the budding yeast *Saccharomyces cerevisiae*. We studied mutational effects in a collection of >80 mutation accumulation strains to capture a snapshot of the effects of mutations that have not yet been filtered by selection. Previous work has demonstrated that these strains contain ~300 mutations, primarily single-nucleotide mutations. By combining genotype information with high-precision microcolony-based growth rate measurements from each strain, we are able to infer estimates for the mean mutational effect size, the shape of the mutational effect distribution, and the proportions of mutations with a neutral, positive, and negative effect on growth rate. We find that mutational effects on growth rate are overwhelmingly negative and highly skewed towards very small effect sizes. Interestingly, our modeling suggests that a single distribution of mutational effects does not account well for the observed distribution of growth phenotypes. Rather, growth rate effects of mutations fall into two categories: single-nucleotide substitutions, which on average have larger effects, and frequent, smaller-effect mutations of unknown origin. The estimated high frequency of mutations in the latter category, combined with the fact that they were undetected by conventional sequencing analysis, point to microsatellite mutations as likely candidates underlying the non-single nucleotide mutational effects. We are performing follow-up experiments to directly assay the effects of *de novo* microsatellite mutations on growth. Our work reveals the spectrum of effect sizes of the mutations on which evolution acts, and points to a likely key role of microsatellite mutations in shaping natural diversity.

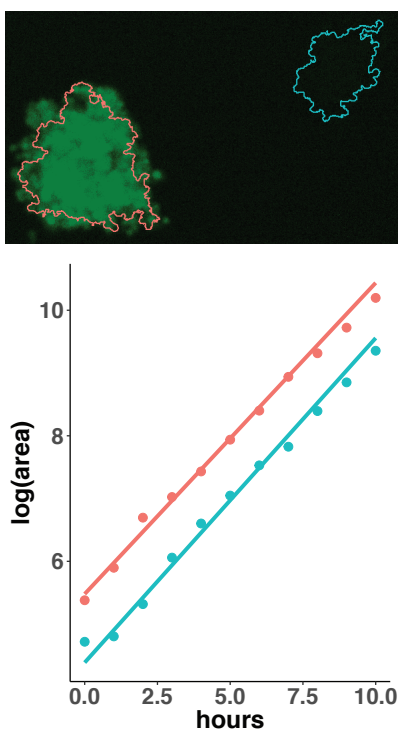
MUTATION ACCUMULATION



Mutation accumulation was performed on diploids with frequent single-cell bottlenecks, thus minimizing the strength of selection. Color selection was used to avoid propagating *petite* mutants (Hall *et al., Genet Res* 2008).
After ~2000 generations, mutation accumulation strains were sequenced, revealing ~8 single nucleotide mutations (SNMs) per diploid MA strain, with a small number of indels and aneuploidies (Zhu *et al., PNAS* 2014)

AN AUTOMATED, IMAGE ANALYSIS-BASED MICROCOLONY GROWTH RATE ASSAY



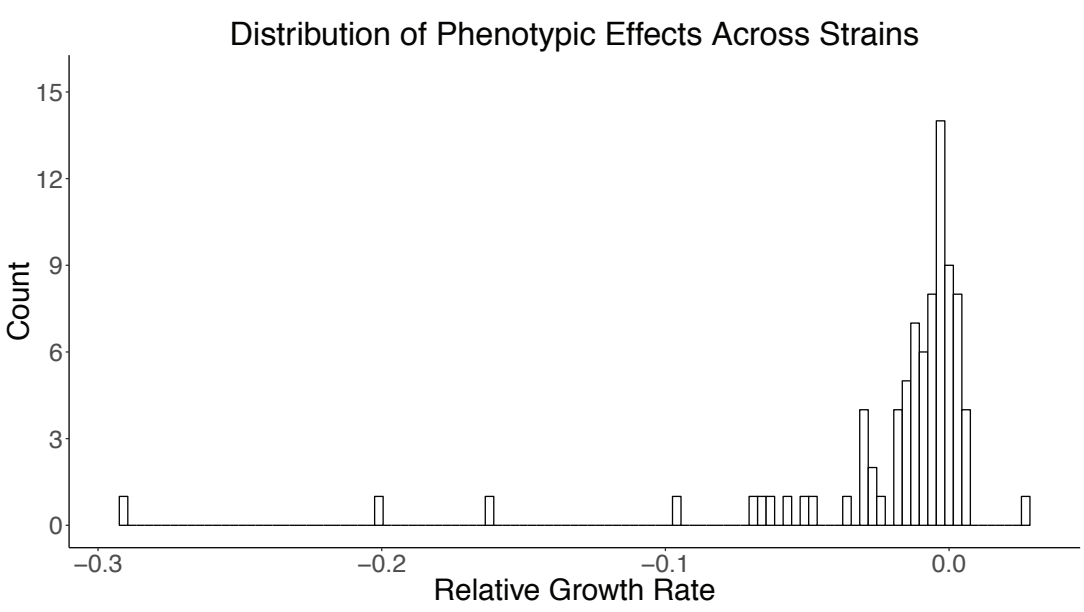


High-throughput microcolony growth rates are measured using automated microscopy and image analysis. Cells from haploid MA strains are attached to the surface of the glass-bottom plate and imaged hourly over the course of 10-20 hours. Colony outlines are automatically detected using custom image processing code. The growth rate is measured as the change in log(colony area) per hour.

Fluorescent imaging at the end of the growth period is used to differentiate between the MA strain and the GFP-marked ancestor strain grown in each well as a reference.

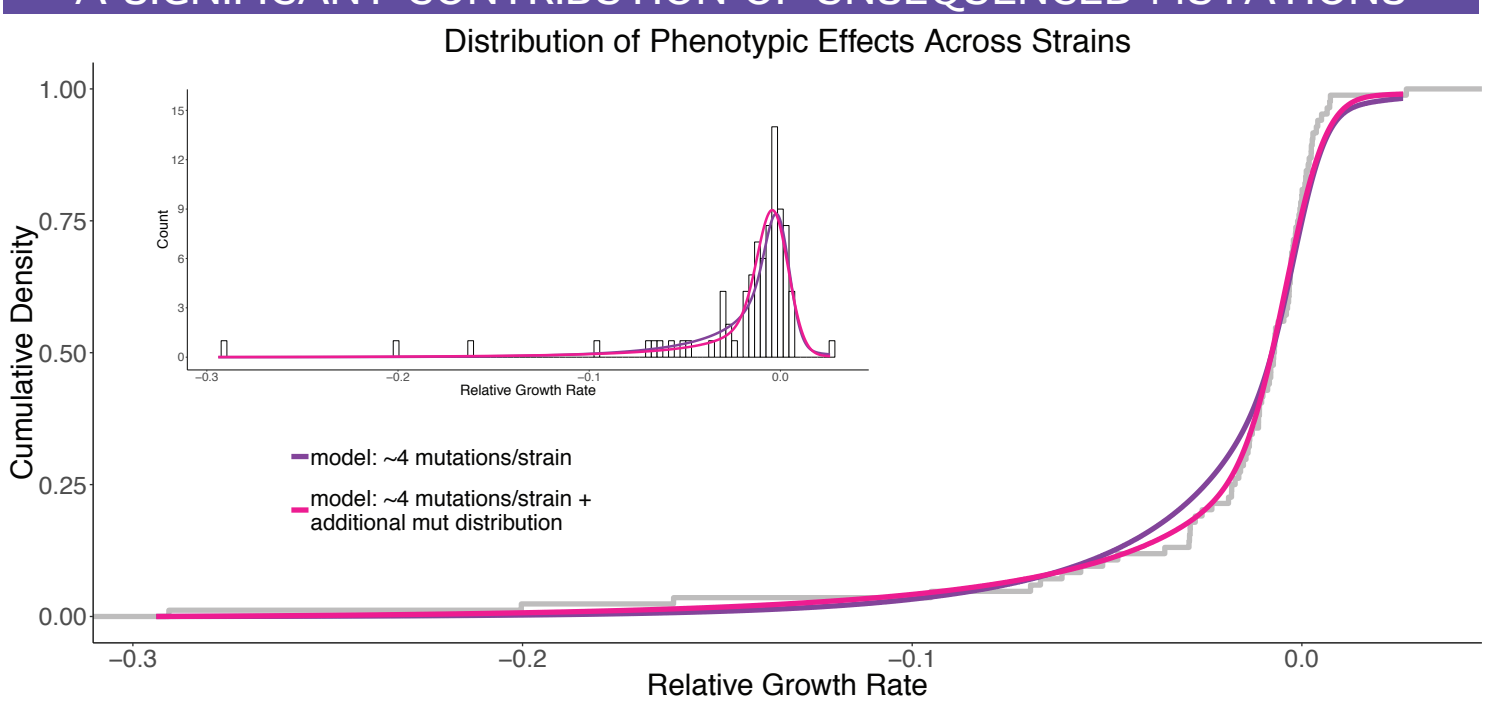
(Plavskin*, Li*, *et al. bioRxiv* 2018)

NET MUTATIONAL EFFECTS ARE PREDOMINANTLY NEGATIVE AND SMALL



The distribution of mutation accumulation strain net effects on growth rate (log of the growth rate of each MA strain relative to the ancestor) show that most strains have a small negative net effect on growth rate; only one strain had a significant positive net mutational effect.

MODELING THE DISTRIBUTION OF MUTATIONAL EFFECTS SUGGESTS A SIGNIFICANT CONTRIBUTION OF UNSEQUENCED MUTATIONS



A model that includes a class of 'unsequenced' mutations (ones that may not be identified by conventional sequencing analysis) provides a significantly better fit ($p < 0.01$) to the relative growth rates of the mutation accumulation lines than a model that includes only a single distribution that accounts for the sequenced SNMs in these lines. We suspect this is due to *de novo* mutations in microsatellites, which occur at a high rate but were excluded from sequencing analysis.

TESTING THE EFFECTS OF SPONTANEOUS MICROSATELLITE MUTATIONS ON YEAST GROWTH RATE

To directly test the effect of microsatellite mutations on yeast growth rate, we are performing a second mutation accumulation experiment in a strain with a deletion of *MSH3*, a component of the mismatch repair complex. *msh3Δ* yeast have a highly elevated microsatellite mutation rate, while retaining a wild type single-nucleotide mutation rate (Haye and Gammie, 2015). By performing a short bout of mutation accumulation in this background, we expect to accumulate many microsatellite mutations but few or no single nucleotide substitutions.
By estimating the mutational effect distribution of microsatellite mutations in isolation, in combination with identification of such mutations in our original strains through high-depth sequencing, we hope to disentangle the relative contributions of different mutational types to a complex trait in yeast.