Mutation spectrum variation in *Saccharomyces cerevisiae* and beyond

@pyjiang

pyjiang@uw.edu

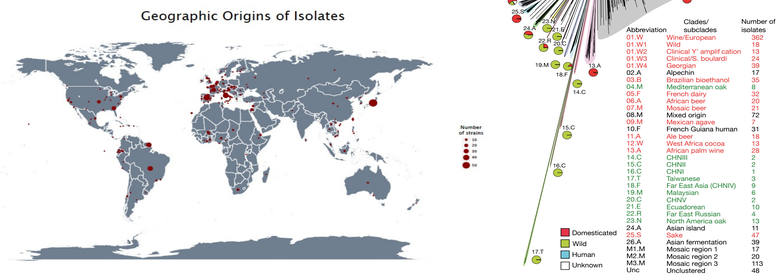
Pengyao Jiang[1], Anja Ollodart[1,2], Alan J. Herr[3], Maitreya Dunham[1], Kelley Harris[1]

[1] Department of Genome Sciences, University of Washington,[2]Molecular & Cellular Biology in Seattle,[3]Department of Pathology, University of Washington

Introduction

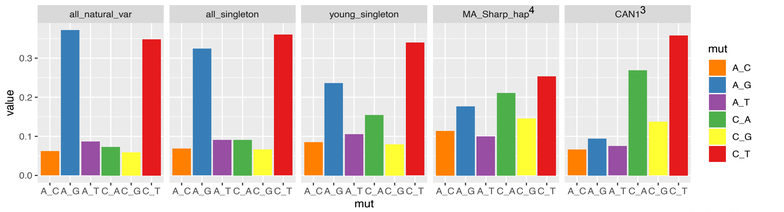
Mutation, the ultimate source molecular variation, is a process can be affected by both genetic variation in DNA repair genes and environmental stressors. Natural genetic variations can be seen as “fossils” of past mutations that have also gone through any natural selection and genetic drift the population has experienced. Differences between populations in mutation spectrum, i.e. the relative frequencies of different types of mutations, can indicate that mutational processes themselves have been evolving.

Recently, the 1002 Yeast Gnomes project surveyed 1011 strains from natural populations of *Saccharomyces cerevisiae* in diverse environments across five continents with deep coverage (200x)[1], providing an ideal resource for a mutation analysis that will parallele what has been done in humans[2].



Abbreviation	Clades/ subclades	Number of isolates
01.W	Wine/European	362
01.W1	Wild	18
01.W2	Clinical 'X' amplif cation	13
01.W3	Clinical/S. boulardi	24
01.W4	Georgian	39
02.A	Alpechin	17
03.B	Brazilian bioethanol	35
04.M	Mediterranean oak	8
05.F	French dairy	32
06.A	African beer	20
07.M	Mosaic beer	21
08.M	Mixed origin	72
09.M	Mexican agave	7
10.F	French Guiana human	31
11.A	Ale beer	18
12.W	West Africa cocoa	13
13.A	African palm wine	28
14.C	CHNIII	2
15.C	CHNII	2
16.C	CHNI	1
17.T	Taiwanese	3
18.F	Far East Asia (CHNIV)	9
19.M	Malaysian	6
20.C	CHNV	2
21.E	Ecuadorean	10
22.R	Far East Russian	4
23.N	North America oak	13
24.A	Asian island	11
25.S	Sake	47
26.A	Asian fermentation	39
M1.M	Mosaic region 1	17
M2.M	Mosaic region 2	20
M3.M	Mosaic region 3	113
Unc	Unclustered	48

Mutation Accumulation (MA) studies and natural genetic variation tend to yield different mutation spectrum estimates



n=613,246 n=102,505 n=1735 n=351 n=212

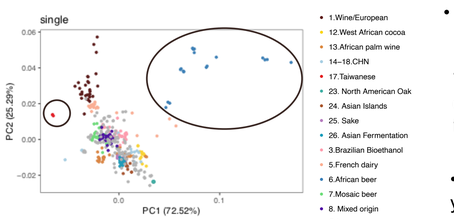
MA Natural

transition / transversion : 0.6-1 3-3.2

pairwise chi-square test of mutation spectrum counts are all statistically significant (p <0.05)

Mutation spectrum differences recapitulate yeast population structure

In this PCA, each individual is summarized by a 6-dimensional vector summarizing the relative proportions of A>C, A>G, A>T, C>A, C>G, and C>T

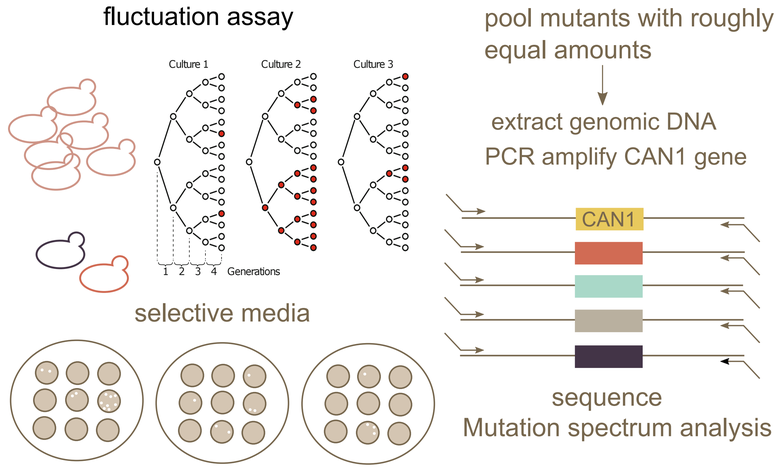


- For each individual, derived alleles are classified into the above 6 types (collapsing strand complements). The first two principal component are calculated from this 6xn mutation frequency matrix.

- African beer yeast a Taiwanese yeast show evidence of distinctive mutational signatures

Experimental design to measure *de novo* mutation spectrum using CAN1

- The CAN1 gene metabolizes canavanine into toxin, so only yeast with loss-of-function mutations in CAN1 can grow on media containing canavanine
- We can ascertain de novo mutations from African beer yeast, Taiwanese yeast, and other haploid strains by plating samples on canavanine media and sequencing the CAN1 genes of mutant colonies that grow
- Goal: compare mutation spectra of yeast strains under controlled experimental conditions to separate genetic from environmental effects



- Haploid starins from different populations, with one senstive CAN1. Include Taiwanese strains and African beer strains. (needs construction)

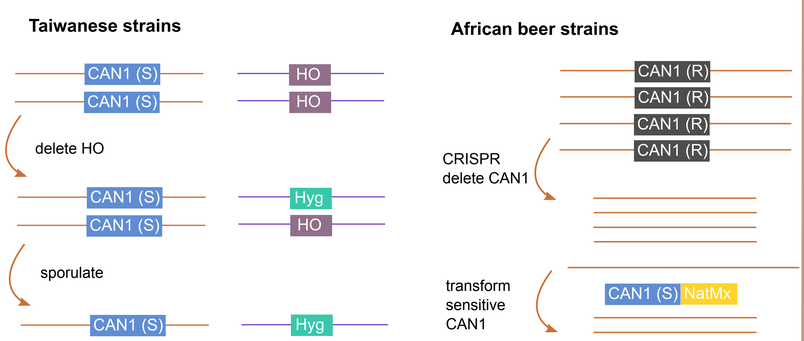
- 1) Quantify mutation rates using flucuation assay.
- 2) Quantify mutation spectrum after pooled sequencing.

- Aim to collect ~300 mutants from each strain.

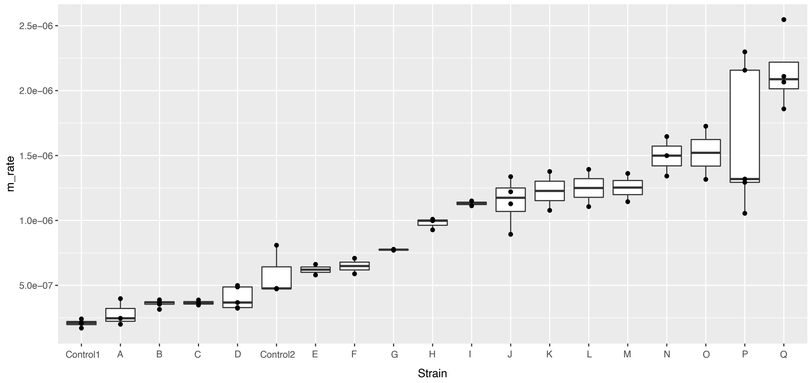
- Each pool contains ~35 mutants. (each mutant frequency is high enough to be distinguished from sequencing errors (~1%).

- Mutant calling pipeline has been validated by sanger seuqnecing.

Constructing Tawainese and African beer strains with one copy of sensitive CAN1

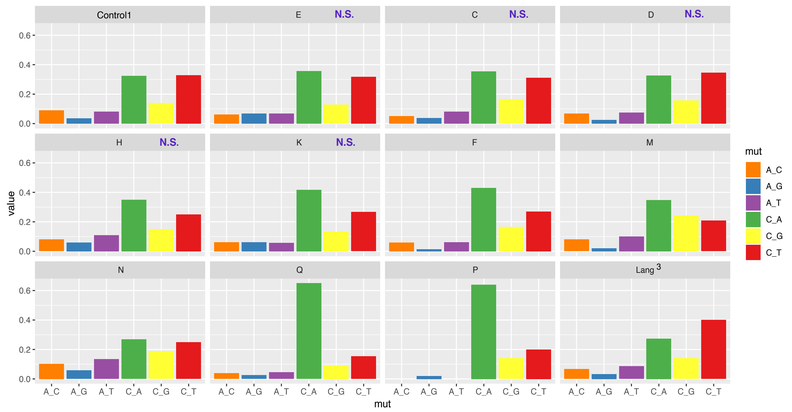


Mutation rate variation across haploid strains



Control1 and Control2 are lab strains that have been used in previous mutation accumulation experiments. A-Q are haploid strains from 1002 Yeast Genomes. Mutation rate is quantified on CAN1 locus.

de novo mutation spectrum from CAN1 in haploid strains



Pooled mutations from each strain are sequenced and called under our pipeline. Mutation spectra are calculated for each individual haploid strain. Chi-sqaure tests are performed from each strain to the Control1 strain. Strains that do not show a significance in mutation spectrum differences are marked "NS" after each strain's ID. Note that strains Q and P which show the excessive C->A mutations also exhibit the highest mutation rate in our assay.

Reference

1. Peter J, et al. 2018. Genome evolution across 1,011 Saccharomyces cerevisiae isolates. Nature, 556(7701), p.339
2. Harris, K. and Pritchard, J.K., 2017. Rapid evolution of the human mutation spectrum. Elife, 6, p.e24284.
3. Lynch M, et al. 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proceedings of the National Academy of Sciences, 105(27), pp.9272-9277.
4. Sharp NP, et al. 2018. The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. Proceedings of the National Academy of Sciences, 115(22), E5046-E5055.
5. Zhu YO, et al. 2017. Extremely rare polymorphisms in Saccharomyces cerevisiae allow inference of the mutational spectrum. PLoS genetics, 13(1), p.e1006455.
6. Lang, G.I. and Murray, A.W., 2008. Estimating the per-base-pair mutation rate in the yeast Saccharomyces

This work is supprted by Burroughs Wellcome Fund career award at the scientific interface to Kelley Harris.