**Genotyping Project**

Experiments to do:

2) Redo the DNA calibration experiment.

3) Design a restriction digest/PCR assay to figure out if the jagged clonal interference peaks are real or not. Definitely do the 12.5K timepoint.

4) Test MAGE by doing allelic replacements of nadR..2.snp onto the losing clade. Is there epistasis, or is the effect additive?

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Rich’s email:

But to use time most efficiently, and given complexity of the problem, I suggest you work up a written outline of:  
  
What are your questions and/or hypotheses?  
  
What do you think the data you already have collected tell you about the answers?  
  
Provide a couple of proposed alternative "designs" for the remaining plates.  
  
What do you think each design can, or can't, say to clarify the answers?  In other words, provide pros and cons for your alternative approaches.  
  
I often suggest trying to draft a title, abstract, and sketches of draft/hypothetical figures for a paper at this stage of a project, as that exercise will often "tell you" what needs to be done to produce and/or wrap-up a story, and to convince readers and committee members.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Under clonal interference, relative selection coefficients are lower, while the beneficial mutation supply rate remains the same. So genetic diversity increases.

We see some classic sweeps, some peaks (clonal interference), and one region of jagged peaks before the fixation of 10 alleles.

We find that sometimes multiple mutations fix simultaneously.

One example of this is a sweep that fixes infB, malT, mrdA, and nagC between 4500 and 5000 generations.

The next fixation occurs sometime between 13000 and 13500 generations, when araJ, gltB-del-d, hslU, nadR..2-snp, nagC, pflC-del-d, rpsM, yedW-yedX-snp, yghJ-snp, and yhdG-fis-snp fix simultaneously. Ten mutations fix! Of course, this is due to clonal interference.

topA and spoT sweep together by 1500 generations, then yegI sweeps by 2500, and ybaL-ins-d by 3000 gens. So it’s not clear if mutations tend to fix sequentially or in groups. Presumably, clumps of mutations sweep together whenever there’s a clonally interfering clade.

Between 7000 and 12000 generations, acs-nrfA-snp is interfering. rpsA-snp has a similar but not identical pattern to acs-nrfA-snp’s interference. atoS-snp interference peaks at 9500 generations. elaD-snp peaks at 10500 generations. hsdM-snp peaks at 10000 generations. maeB-talA peaks at 10000 generations. mreB-snp peaks at 2000 generations. nuoG-snp peaks at 11000 generations. nuoM-snp peaks at 10000 generations. pykF-snp has a very small peak at 1500 generations. pykF…2-ins has a peak at 1500 generations. pykF..3-ins has a very small peak at 1500 generations. pykF..4-snp has a small peak at 2000 generations. rspA-ynfA-snp has a small peak at 15000 generations. yijC-snp peaks at 1000 generations.

The previous inferences come from GenotypePredictor.R and the ‘movie graphs’ I made for lab meeting. But neither are really ready yet for presentation, or for making further conclusions on sweeps.

**I need to graph clonal interference peaks in a separate “movie graph” to compare with periods when multiple mutations fix.**

Question: do multiple mutations fix at the same time as a consequence of clonal interference? What is the proper model to test (covariation between number of mutations fixing, and the size/length of a competing variant?)

Question: what kind of resolution does the clone data give us in terms of the order of mutations?

Answer: I’m not sure. I’ll have to investigate the tree that Jeff made in greater detail.

Clonal interference slows down the pace of fixation, but not the mutation supply rate. So, longer and more diverse trajectories are explored before the fixation of one of these trajectories.

Idea for further work: Can epistatic interactions between genes predict the order in which mutations occur in an experimental population? First, use MAGE to generate all combinations of mutations for a sweep containing multiple mutations. Then, do fitness assays across the library.

Subidea: make a device that does high-throughput fitness assays?

Then, do colony PCR from samples taken from the 10K and earlier time points, to find genotypes with different combinations of the alleles of interest. (An undergraduate could be really useful for this last bit.). Find ‘nested’ genotypes to infer the order of mutations.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Genotyping Experiments:

3a) Design a restriction digest/PCR assay to figure out if the clonal interference peaks and valleys are real.

3b) Definitely do the 12.5K timepoint.

4) Test MAGE by doing allelic replacements of nadR..2 snp onto the loser. Is there epistasis? First sequence some losing nadR genes to see if there are unassayed mutations already there.