7.344 Directed Evolution: Engineering Biocatalysts
Spring 2008

Session 12 Lecture Notes

1. The authors have developed a new strategy to link genotype and phenotype based on compartmentalization. Essentially is like ribosome display but prevents the crosstalk we identified as a problem last week. A protein and the gene that encodes it are encapsulated within an emulsion along with substrate that is somehow detectable. The key in this first paper is that the gene is the substrate. How do the authors prevent crosstalk here before emulsification? Is this reasonable? Where is the data?!

2. The compartments are made by adding an aqueous reaction mixture to a stirring solution of mineral oil containing surfactants. Droplets are 2.6 um in diameter – about the size of bacterial cells. The volume of a droplet is about 5 femtoliters. One mL of the emulsion contains about 10^{10} compartments (or genes). How are the emulsions broken up – ether? Does everyone understand this? Extraction.

3. The enzyme used for this first paper is the DNA methyltransferase HaeIII. The authors select for methylated DNA based on its inertness to digestion with HaeIII endonuclease. Those that make it through are used in subsequent rounds of reactivity and selection. Can easily see this using agarose gel.

4. Lots of good controls: no gene retrieval when 1:1000 mix is PCR amplified, or not emulsified, or without IVT reaction mixture, or if cleaved with an alternate restriction enzyme following the reaction. The yield isn't, however, 100%. Why is this? Maybe the digestion isn't complete, but what about crosstalk? Also proved that enrichment of genes as observed by gel correlates with activity increase.

5. Take a look at Figure 4 Panel b. Good mixing experiments presented here. Looks like the 1st round enrichment is about 1000-fold when the 1:1000 ratio is used. Tails off in first round with more inactive enzymes.

6. Biggest pitfall: Hard to do this with other enzyme classes as presented since DNA must be linked to the substrate. How could you do this? Doesn't evaluate kinetics. They propose that compartmentalization will allow for reactivity wit substrates that aren't attached. How does this work? Seems like a lot of engineering to me.

7. The second generation of the method uses two separate compartmentalization steps. Walk through figure 1. This is necessary to make compartments that contain only one kind of enzyme and also enables examination of kinetics – both the use of beads and the use of the second compartment to normalize reaction start times help with this.

8. Take a good look at Figure 5. The authors prove to themselves that enrichment requires compartmentalization. Make sure everyone understands chemistry.
9. Saturation mutagenesis is used at particular sites. Library is compartmentalized in to 4 sublibraries, why? Very slick method of amplifying for the next round of selection using primers that are already biotinylated.
10. The authors find a mutant that is 63 times more active that the wildtype enzyme. Total mix has 44% of wild type activity after 6 rounds – is this comparable to previous methods? Maybe a little bit poor even though enrichment seems quite good. What if library generation is bad…Also, the reaction is run for 16hrs, possibly way too long to get meaningful kinetic data. The enzyme is so active you are likely at steady state very quickly. So now poor catalysts have an equal chance for retrieval as compared to good ones. (look at figure 8)
11. Seems like a lot of engineering and tethering to me – is this strategy really better than an in vivo system? One plus – affinity selection using anti-product antibodies will allow for screening of much bigger libraries by doing a selection first. Why is this important? Because FACS is slow!! They say that selection is performed for a soluble substrate and is for turnover – it is true that they can assay kinetics, but is the substrate really soluble? Or is it just that this enzyme is so fast (2280 s-1) that the substrate reacts before it gets sequestered? This could be a big problem in the evolution of enzymes with little activity.