Part A:

1. Top 10 unique probes for red channel:
   1. DET 0748
   2. DET 0078
   3. DET 1437
   4. DET 1568
   5. DET 1427
   6. DET 0079
   7. DET 1407
   8. DET 1436
   9. DET 1428
   10. DET 1206
2. Top 10 unique probes for green channel:
   1. DET 1296
   2. DET 0748
   3. DET 0079
   4. DET 1437
   5. DET 0078
   6. DET 1568
   7. DET 1407
   8. DET 1569
   9. DET 1438
   10. DET 0381

There is some overlap between these two lists. This suggests that the overlapping probes would not be very useful for distinguishing between the two samples of cDNA, but that they bound easily with both. It could suggest that those particular probes aren’t very specific.

1. Sort by log ratio between the channels. Top 10 unique:
   1. DET 1243
   2. DET 0284
   3. DET 1477
   4. DET 1511
   5. DET 0420
   6. DET 0313
   7. DET 0010
   8. DET 1278
   9. DET 1554
   10. DET 0558

This indicates high gene expression because the ratio is larger. The negative values indicate low gene expression because the probe isn’t binding with the PCE present.

1. The Control vs Control plot is more tightly linear because there is no other gene expression affecting the fluorescence.
2. Volcano plots!

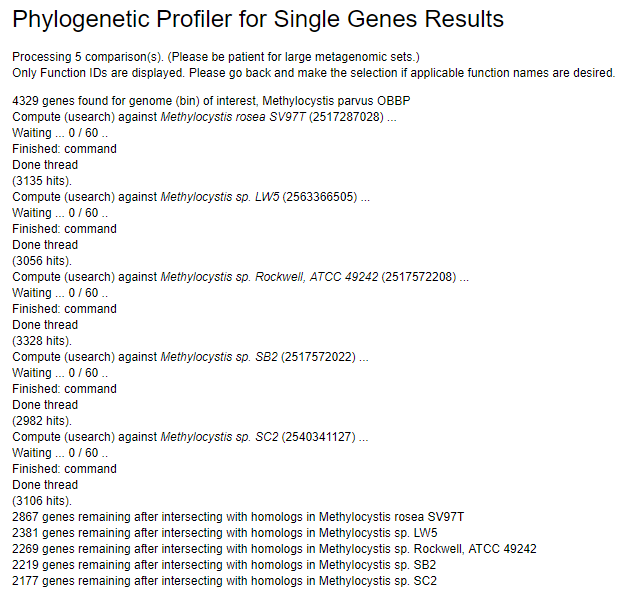
These plots do look different!

Firstly, there are many more points on the PCE-control graph than the Control-Control plot. This would indicate that more samples are determined to be different depending on the growth media, as would generally be expected if the data is to show anything. Another difference is the spread. The Control-Control plot shows a tail to the left suggesting some of the spots were far more fluorescent in some spots. This seems like it could indicate and error in those points.

1. Hacking Excel for a heatmap!
   1. From the ‘heatmap’ it does seem that the PCE-control column does have more extreme values than the control-control column. I used red as high and green as low, and the control-control column tended towards more green and the PCE-control column tended towards yellows/reds. This does fit with the volcano plots and X-Y plots previously generated because it also shows a greater difference between the PCE and the control through the expression of certain genes.
2. 3 genes with differential expression (p < 0.01)
   1. DET 0553: ATP synthase F0, A subunit( EC:3.6.3.14 ) (atpB)
      1. [GO:0015078](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0015078) - hydrogen ion transmembrane transporter activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0046933](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0046933) - proton-transporting ATP synthase activity, rotational mechanism [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]
      2. Involved in oxidative phosphorylation and photosynthesis
      3. Genomes with homolog: other dehalococcoides, chloroflexi bacterium.
   2. DET 0751: Ribosomal Protein L35 (p-value = .00978) (rpml)
      1. [GO:0003735](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0003735) - structural constituent of ribosome [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]
      2. Involved in : Ribosome
      3. Many genomes with homologs: *mycobacterium marinum, heliobacter pylori…*
   3. DET 1158: nitrogenase iron protein( EC:1.18.6.1 ) (p-value =.00987) (nifH)
      1. [GO:0051539](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0051539) - 4 iron, 4 sulfur cluster binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0005524](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0005524) - ATP binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0005524](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0005524) - ATP binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0018697](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0018697) - carbonyl sulfide nitrogenase activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0051536](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0051536) - iron-sulfur cluster binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0046872](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0046872) - metal ion binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0016163](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0016163) - nitrogenase activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0016163](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0016163) - nitrogenase activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0000166](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0000166) - nucleotide binding [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0016491](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0016491) - oxidoreductase activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]  
         [GO:0016491](http://www.ebi.ac.uk/ego/DisplayGoTerm?id=GO:0016491) - oxidoreductase activity [evidence=[IEA](http://www.geneontology.org/GO.evidence.shtml)]
      2. [Chloroalkane and chloroalkene degradation](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=KeggMap&page=keggMap&map_id=map00625&gene_oid=637120721&myimg=0)  
         [Nitrogen metabolism](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=KeggMap&page=keggMap&map_id=map00910&gene_oid=637120721&myimg=0)  
         [Microbial metabolism in diverse environments](https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=KeggMap&page=keggMap&map_id=map01120&gene_oid=637120721&myimg=0)
      3. Other dehalococcoides, achromatiums, and many others!

Part B

1. Core genome determination 2177 genes are conserved, ~half



Once all the genomes have been compared, 2177 genes are conserved.

1. Can everyone get phb starting from acetyl-CoA
   1. Yep as far as the database knows, the blue enzymes are enzymes present in all organisms.

