Iddo/Scott phone meeting – morning Feb 16.

* Scott showed Iddo
  + Multiple testing correction of gene expression value (EV) prediction
    - EV = 1 + 1\_[BF] + %Actino + %Firmicuties
    - There may still be some statistically interesting results
    - That will be interesting if there is some biological interpretation available for the significant findings
* The microarray data (54k probes -> qc reduced to -> 16k -> with 8k OGS annotated) was discussed
  + The possibility of Iddo doing better annotation using the probe sequences was raised. Iddo would use interproscan on the computing cluster in ohio
* Scott sent Iddo plots of the metabolic function breakdown across samples for all SEED categorizations.
  + Metabolic function breakdown with Virulence and several other categories (and subcategories) appeared interesting
  + Iddo is going to look closer at the data to find areas worth further consideration
  + Is the immune system of 3 month old infant even active?
* We began discussing correlating individual genes (EVs) with metabolic profiles
  + Scott showed Iddo example results for 10 genes with smallest q-values versus the top 10 SEED level 1 metabolic profiles
  + There was not much by way of interesting results
  + Iddo suggested we pursue a similar analysis using the virulence category

NEXT STEPS

* Scott resends colored metabolic profile breakdown plots to iddo.
* Iddo examines plots for to find areas of further interest
* Scott produces a more complete immunology host gene list
* Iddo discuss immunology implications of this list with his in-house immunologist – also, are 3-month old babies immune systems active?
* Scott plots first and second pc’s of SEED level 1 pca
* Scott examines biplots as a way to see what PCs are.

Iddo provided Scott with

* GO review paper
* PCA chapter
* Unifrac website… industry standard for Phylum breakdown

Scott plans to take a look at a multivariate analysis book