I. Project Start Date: *Oct 9 ,2013*

II. Project leader: *Alex Belieav*

III. Data analysis leader: *Bryan Linggi*

III. Location of project files: *afp://we25473//Users/ling551/Documents/R\_onAir/Belieav\_Axenic\_10092013*

Experimental Design

V. Experiment Information

a. Experiment type:

New Data Analyses: generated in house, but not “new”

Public Data Analyses

Gene Expression

Microarray

RNA-Seq

RT-PCR

Other:

Epigenetic analysis

Methyl-Seq

ChIP

Protein expression

Mass-spectrometry

Other:

b. Experiment layout

Experimental Hypothesis: *Differences in light and oxygen will dramatically alter the level of gene expression of pathways that contribute to the cell survival. \* Alex says he will provide specific hypotheses about the experiment soon (10/9/13)*

Output result (check all required):

Powerpoint presentation of highlights

Excel table of raw data

Excel table of normalized data

Excel table with p-values and fold changes

p-value cutoff:

multiple hypothesis correction? y/n

fold change cutoff:

Graph of data summary

Details:

All analyses files (will be archived even if not returned to user)

other: figures for paper,with legends and materials and methods. Goal is to have draft by end of October

Cell Type(s): Synechococcus sp. PCC 7002

References: recent Bryant Lab paper: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3468840/

Conditions:

*1. Synechococuss 7002 OG1 + LK1 High light/ high O2*

*2. Synechococuss 7002 OG1 + LK1 low light/ high O2*

*3. Synechococuss 7002 OG1 + LK1 low light/ low O2*

*4. Synechococuss 7002 OG1 + LK1 high light/ low O2*

*\*all 4 above were run on SOLiD in August 2013. Previously, data was generated by Covance (Illumina), that had 5 samples, adding 1 that had very high light. Only analyzing these 4 here*

Controls \**waiting on input from Alex about this.*

Positive: *high light, low O2 vs low light, low O2*

Expected result: *increase in the expression of photosystem II genes (not sure about this)*

Negative: *low light, low O2 vs low light, high O2*

Expected result: *no change in photosystem II genes (not sure about this)*

c. Power analysis

Replicates needed for desired power: *experiment already performed \**

Type of replicate: *no replicate. This is a single culture from chemostat. Illumina samples (mentioned above) are from same chemostat run (different lot of cell pulled , all are at steady state by definition of chemostat) and may be referenced for comparison.*

1. VI. Preprocessing Steps:

(Mod from Faraway 2002)

Set up Git locally and on GitHub. Record details of Github locally): *afp://we25473/Documents/R\_onAir/EGF\_gage analysis*

Github address: *https://github.com/bedward1/EGF\_gage.git*

Data collection

how are data collected (random sampling?): *affymetric microarray. Sample prep randomization methods unknown (were preformed other investigators before I arrived.)*

Run preprocessing code in R

Output is

1) html file in

*we25473/Documents/R\_onAir/EGF\_gage analysis/report/preproc\_report.html*

2)R.data for input to next steps

*we25473/Documents/R\_onAir/EGF\_gage analysis/diagnostics/Preproc.R.data*

1. Initial analysis
   1. Correlation calculations and plots
      1. Save Rdata for next steps in :
         1. *afp://we25473/Documents/R\_onAir/EGF\_gage analysis/tests/*
      2. Save html report in
         1. *afp://we25473/Documents/R\_onAir/EGF\_gage analysis/reports/*

3) Pathway analysis

a) perform Gage analysis

use Kegg pathways

1. output table to /tests folder

comparing gene pathways at different densities and inhibitor presence

note changes in pathways and level in pathways.