I. Project Start Date: *Oct 7 ,2013*

II. Project leader: *Bryan Linggi*

III. Data analysis leader:*Bryan Linggi*

III. Location of project files: *afp://we25473/Documents/R\_onAir/EGF\_gage analysis*

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: *EGF affects different cell processes depending on concentration of ligand*

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)

Cell Type(s): *human mammary epithelial cells 184 A1 (hmec)*

Conditions:  *EGF at concentrations 0 to 100ng/ml*

*Inhibitors, mek, U0126*

Controls

Positive: *EGF at 1hr*

Expected result: *dusp6 expression increases*

Negative: *no EGF, 4 hr*

Expected result: *no change in dusp6 expression*

c. Power analysis

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

Type of replicate: *duplicate or triplicate*

1. VI. Preprocessing Steps:

(Mod from Faraway 2002)

Set up Git locally and on GitHub. Record details of Github locally):

Github address: https://github.com/bedward1/

Data collection (export to 'diagnostics' folder)

how are data collected (random sampling?):

is there non-response?

are there missing values?

how are data coded?

how are qualitative variables represented?

Do I need to convert them to factors?

yes

no

if yes, describe steps

What are units of measurement?

Can I get rid of extraneous digits to save space?

yes

no

If yes, how is it done?

4. Initial data Analysis (export to 'diagnostics' folder, see location from )

a. Print out summaries of data

look for coding errors (NA converted to 0s)

b. Make box plots and examine variance

c. Make scatter plots for correlation analysis

3.

check for missing values

use boxplot or summaries (See Farraway PRA pdf)

4. recode for all factors

5. perform anova

6. output to single file for different primers