Lesson Plan Day 1

1. Introductions of ourselves and of the students. Provide name plates.
2. Big intro- Colin
    a. Why a data expedition? Why do we care about analyzing complex data? Working with unpublished data.
    b. List the overall goals for the module
        i. Learn to visualize and interpret complex data
        ii. Genetic variation -> gene expression -> phenotypic changes
        iii. Not be scared of programming!
    c. Go over plan for today
3. Biofilm and complex colony morphology intro- Colin
    a. Key concepts: important in natural settings also have direct implications for humans. Caused by excretions by the yeast.
    b. Nutrient dependent.
    c. Adhesion in complex colonies is stronger.
    d. Have to have growth in the top part of the cells despite nutrient limitation
4. Intro to R- Both
    a. Get them started opening up the programs, get them oriented to the Console and how to run lines of code. Set them loose with the workbook.
    b. Provided challenge questions that entailed them entering data into an object and graphing it via novel functions hist and density.
    c. In the future include more exploration of options and more challenge questions. Include an earlier explanation of how to get help.
    d. Key concepts: objects, functions, arguments in functions, plotting
5. Bulk segregant analysis to figure out genetic basis of complex traits
    a. Powerpoint intro to complex traits -Colin
        i. Concept of additive allelic effects
        ii. Concept of segregation of allelic variation evenly into offspring
        iii. Alleles that influence a trait are arranged spatially on chromosomes and most alleles don't have any effect
    b. Bulk Segregant activity on Digital organism -Liana
        i. Powerpoint
            1. One chromosome, 20 loci that are heterozygous, arbitrarily defined to be A and B alleles, Created 24 offspring.
            2. Add in slide that puts the creation of the segregants in the same context as above.
            3. Add in slide that shows that there are 50% of each allele in the offspring population.
            4. Show an example of how you figure out the "phenotype" of the organism from A and B alleles that matter. Show that both A and B alleles can alter the complexity phenotype.

ii. Create histogram of the "phenotype" of the twenty strains on the board
iii. Have each group "sequence" one bulk- add up the frequency of A and B alleles and calculate the absolute deviation. Have each group write up their results on the board. Then have them start on the work sheet where they graph the difference between the two bulks and then answer all the questions. Then discuss as a class.

c. Take home points about peaks
   i. Big allelic effects are easier to find than small allelic effects
   ii. Effect of linked loci and random variation on peaks
   iii. Connect to Granek Figure 2 explicitly.

6. Lecture
   a. Interpret the Granek BSA figure (projected on screen) what genes are associated
   b. Introduce PKA
   c. Introduce cAMP and adhesion in the simple/complex
7. ~~Exploring the gene expression data for the first time- R workbook-~~
8. Initial exploration of the physiological parameters of CCM (cAMP, adhesion)
   a. Importing data to R
   b. Association of FLO11c with adhesion (written out)
   c. Association of CYR1c with cAMP (make them figure out the code).
   d. Association of HOT1c with cAMP
9. Wrap-up
   a. Go over which parts of the goals we accomplished that day
      i. Not being afraid of R!
      ii. Figuring out the connection between genetic variation and phenotype.
   b. Highlight where we are going next-
      i. What genes were found to alter complex colony morphology? What do they do? How do they influence gene expression
      ii. How can we make these massive datasets you just brought into R simpler?
      iii. Understanding heatmaps!!!

Lesson Plan Day 2:

1. Reboot of Day 1
    a. First three worksheet for the day to get people back in gear.
    b. Tie in what we did to the larger goals
        i. Visualize and interpret complex data
        ii. Genotype-> Gene Expression -> Phenotype
        iii. Don't be afraid of programming in R!
    c. Where are we going today?
        i. intermediate role that gene expression plays in creating phenotypes.
        ii. Hierarchical clustering
        iii. Interpreting heatmaps
2. Recap of biology
    a. BSA, PKA, cAMP, adhesion
    b. Time out for the central dogma explanation and.
        i. Genetics is not magic!
        ii. Question 2-What are various ways of changing gene exp.
    c. Question 3-What type of allele is cyr1
        i. Provide background and walk through the content of figure 3 (segregants, over-expression lines, empty line)
        ii. Have them answer Question 3
3. Dendogram/Hierarchical clustering
        i. Intro + example dendrogram on board
        ii. Intro to correlation
        iii. Worksheet (catch them when looking at graphs and make sure that they understand axis)
        iv. Correlation to heat map and anti-correlation
        v. Ask Question 1 on gene exp portion…
4. Interpreting sample and gene clusters at two data set of entire subset
        i. Talk about the measures themselves. Changes in expression and not absolute. Everything is relative to the average.
        ii. AB and 12
        iii. ABC and 123
        iv. Is this consistent with our predictions from Question 4?
        v. Kinds of genes as well as number
5. Introduce results from genotyping:
        i. A way to see how consistently a genetic variant associated with the phenotype segregates with the phenotype. How do these results differ from BSA? Draw attention to CYR1 and HOT1
        ii. Take home- lots of genes influence complex colony morphology
6. What are these "gene effects"?
    a. Emphasize that variants can effect either the expression or the promoter function or on something else entirely
    b. Draw out exactly how variation can effect phenotype

7. Exploring the effect of alleles on gene expression
    a. Effect of FLO11 expression on adhesion
    b. The effect of the FLO11c allele on adhesion
    c. Effect of HOT1 allele on CYR1 expression
8. Wrap-up
        i. Revisit course goals
        ii. Feedback from students

Day 3:

1. Recap of Day 2
    a. Recap original biological background (dextrose and Ras-PKA)
    b. Present switch from aerobic to anaerobic processes
    c. Show heat map
2. Plotting allelic variation on the gene expression heat map
    a. Workbook 3.4
    b. Discuss three questions
    c. Recap the genotyping data
    d. Redraw the line graphs of the expression of the clusters in the different samples
3. Physiological consequences/Functional enrichment
    a. Present the problem of "what are the genes"
    b. Introduce concept of a gene ontology having genes -> terms
    c. Tell them that we picked three representative genes.
    d. Fill out the board with what the genes represent and what the expression of the cluster was
    e. Segue we have three genes in a related process, but how likely is that to have happened by chance?
    f. Pretend there's a cluster of 10 genes. How likely are you to have gotten 7 that are "mitochondrial." Do this as an activity with M&Ms.
    g. Show on the computer that you get the same distribution as with the M&Ms.
    h. Make a simpler, more intuitive function for the computer based sampling based on the 'sample()' function.
    i. Pass out cards that list the 1) cluster size, 2) the total amount of each functional category in the whole genome, 3) the amount of each category in their cluster. Ask them to determine if any are enriched/significant. Make them do the graphs.
    j. Look at the GO terms, make them add some more subtlety to the board. Go around and have them talk about it.
4. Big Picture Wrap UP
    a. Think, pair, share, of big picture conclusions from the analysis so far
        i. My takeaways are: 1) gene expression differs widely between simple and complex colonies, 2) Different allelic variants are present in complex colonies.
5. Feedback Forms
6. Final project
    a. Ask them to synthesize an answer to a question about FLO11c
    b. Ask more questions they haven't done
    c. Unguided exploration
7. Bring it back together
    a. Draw the plots on the board and talk about what they are and what they mean

NOTES:

1. Limit types of graphs to boxplots, points, densities, (no violins). Avoid log-scale
2.