**Clinical and Research Genomics Assignment #2**

**From Lecture\_04-05 (March 28th):**

**RNA-Sequencing, Epigenomes, Gene Fusions, and DNA Modifications \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Assignment: Answer questions about RNA-Sequencing and their methods**

**Due Date: 10:00AM on April 4th**

This week has two sections: short-answer questions and a data exercise.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

(State whether the statement is true or false, then explain with one sentence).  
1) There are 20 types of RNA present in cells.

2) Once spliced and 5’ capped, an RNA is no longer modified.

3) Gene fusions identified by RNA sequencing always correspond to rearrangements in the DNA as well.

Essay Questions  
1) If you were designing an experiment with RNA-Sequencing for each sample:

a.) How much sequencing is required to assay each sample?

b.) What are the parameters that affect this depth of sequencing?

c.) How would you calculate an expression value?

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Epigenome and DNA Modifications**

Analyze and Contextualize DNA Methylation data from a RRBS experiment

If you do not already have it installed, install the free statistical program R on your computer:

<http://www.r-project.org/>

You can use the slides on this blog to understand the algorithms that we have discussed in class:

<http://zvfak.blogspot.com/2013/03/epiworkshop-2013-dna-methylation.html>

If you have trouble getting the package for R from the tutorial, you can download it here:

<http://code.google.com/p/methylkit/>

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Utilize the tutorial to familiarize yourself with R and the program.

Then, from the test dataset within the package, your assignment is to:

1. Check if there is an indication of PCR bias in the experiment.
2. Calculate # of differentially methylated cytosines (DMCs)
3. Annotate the DMCs with genes and discuss implications of that annotation.
4. Check correlation between samples.
5. Cluster samples to see if replicates cluster together

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Please hand the assignment on the day of the lecture, or email if you cannot attend.**

**For any questions, please contact Alexa McIntyre (**[**abm237@cornell.edu**](mailto:abm237@cornell.edu)**), Ebrahim Afshinnekoo (**[**eba2001@med.cornell.edu**](mailto:eba2001@med.cornell.edu)**), or Professor Mason (**[**chm2042@med.cornell.edu**](mailto:chm2042@med.cornell.edu)**).**