Clinical and Research Genomics Assignment #3
From Lecture_10-11 (April 4th): Genome Re-arrangements and Linked Reads

_____

Assignment: Answer questions on genetic variation and clinical genomics.
Due Date: 10:00AM on April 11th

1. How is the positive strand of a chromosome defined?

2. What is the difference between indels, structural variants (SVs), and copy number variants (CNVs)? Why is this a useful distinction to make for both detection purposes and general biology?

3. Where in terms of chromosome and strand would each of a pair of DNA WGS short reads align if they flanked (but did not span) the breakpoint of a fusion joining the 5' end of gene A on the negative strand of chr 1 to the 3' end of gene B on the positive strand of chr 2?

4. We use long-read nanopore DNA sequencing to confirm the gene fusion in (4). These data comprise thousands of single end ~10 Kbp reads, each a 5' to 3' sequence. We align these reads to the reference genome in a piecewise fashion, so a gene fusion will yield a multi-part alignment. What are the expected alignment patterns (successive chromosomes and strands) for reads that span the fusion in (4)? Answers should be written in 5' to 3' order along the read.

5. Describe a strategy to improve identification of short reads in metagenomic experiments (from class or another source).

6. What assumption underlies the Perfect Phylogeny Model? Briefly comment on the relevance of the model to cancer research.

_____

Please hand in the assignment on the day of the lecture or submit it by email beforehand to your TAs.
For any questions, please contact Alexa McIntyre (abm237@cornell.edu), Ebrahim Afshinnekoo (eba2001@med.cornell.edu), or Professor Mason (chm2042@med.cornell.edu).