Genomics Spring 2018  
Group Assignment 2  
due April 7, 2018  
  
**Part 1 - Regulation**  
This part is VERY open-ended.  
  
Each group gets a recent paper. The paper describes a potential clinical association involving a noncoding SNP. The SNP is more than likely in a regulatory or intronic region.  
  
You will want to analyze the SNP from a genomic and regulatory perspective. In particular, try to use tools we have used to learn as much as you can about how that gene is regulated, focusing primarily, but not necessarily exclusively on the SNP region.  
  
Prepare a report. You can start with **what is already known about the gene/SNP**. That should be a short recap**. Then find out more and report it.** **There should be a diagram of the gene model with the SNP highlighted**. **More than one diagram is probably needed**. Figure legends should be detailed and explanatory.  
  
The structure is up to you. Here are some guidelines. They are guidelines, not required questions. It's to give you an idea of the type of direction this could take. You should be thinking of questions particularly relevant to your gene(s)/SNP.  
  
\* What clinical consequences are associated with this? This should not take more than a couple of sentences. Your focus should be regulatory and genomic.  
  
\* What transcription factors or proteins bind the regulatory region that the SNP affects. How do those proteins regulate other genes?  
  
\* Is there ENCODE data that is relevant? Chromatin modifications, DNAse hypersensitivity?  
  
\* Are there other transcription factors/regulatory regions that are relevant?  
  
\* How common is the SNP? Any geographic significance?  
  
\*What further experiments in the wet lab or dry lab would you propose doing to learn more. Think especially about ChIP-seq and RNA-seq (if relevant) and be specific with hypotheses.  
  
Again, the challenge will be to determine a direction. You may want to pick a leader and assign roles. I think you should spend the first week really discussing how to tackle this. You will probably not figure out exactly how this gene is regulated. You could, however, make some hypotheses and develop some insight that you can share.  
  
You will need some heavy discussion within your group, but please feel free to kick around ideas with other groups. There will be a discussion forum for this.  
  
Your papers are (PMID = PubMed ID):  
  
Group A: PMID: 27817866  
Group B: PMID: 27792999  
Group C: PMID: 27580702  
Group D: PMID: 27776952  
Group E: PMID: 27191258  
  
  
**Part 2 - NGS Processing**  
The list below has links to NGS data from the 1000 Genomes Project.  
  
Groups A & D: ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04022/sequence\_read/SRR790192\_1.filt.fastq.gz  
ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04022/sequence\_read/SRR790192\_2.filt.fastq.gz   
Groups B & E: ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04020/sequence\_read/SRR791475\_1.filt.fastq.gz  
ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04020/sequence\_read/SRR791475\_2.filt.fastq.gz  
Group C: ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04019/sequence\_read/SRR790765\_1.filt.fastq.gz  
ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3/data/HG04019/sequence\_read/SRR790765\_2.filt.fastq.gz  
  
Run the paired-end data through the assembly process using either Bowtie or BWA. Filter the BAM file to limit to the region that your part 1 gene is from. Make your region about 100,000 bp or the size of your gene (if it's larger). Create a VCF file from the filtered BAM file.  
  
-Submit the VCF file  
-List the variants in your gene(s) from part 1  
-Submit an image from IGV (or another browser/viewer if you prefer) of the alignments in the region of your SNP from part 1  
-Provide a brief summary of your results.  
  
Groups  
Here are your group assignments:

|  |  |
| --- | --- |
| Group (Team) | Members |
| 1 | Sai Shruthi, Matthew Margolis, Mike Colgan |
| 2 | Chee Leong Cheng, Arman Seuylemezian, Yohan Sumathipala |
| 3 | Sara Contente, Anil Surathu, Matt Cook, Alexander Mestre |
| 4 | Hadi Issah, Michelle Jimenez, Luis Aparicio |
| 5 | Julie Garcia, Marka Van Blitterswijk, Hannah Weidmaier |

Extra Credit Opportunity: Group Project Management (Optional)

(We recognize actual field works in bioinformatics often involve multi-disciplinary team work. Bioinformaticians (or bioinformaticist) often work together with post-docs and principal investigators bridging the gap and solve the problems they have in scientific research settings. In industry, bioinformaticians (or bioinformaticist) are the people understand the whole picture of the projects so that they often take a role of project manager. In addition, students are passionate and would like to their best job for the group projects but their schedules are pretty tight. We found some coordination and guidance on project management were helpful. Therefore, we decided to provide this opportunity.)

Group Project Management: total 1 extra credit point (Optional)

Your group discuss whether to complete project management extra credit homework or not

if your group chooses not to do it, that is ok. Your group will not be penalized in any way and will not receive the extra credit point.

if your group choose to complete the group project management homework for group project #1, then everyone in your group will receive one extra credit point.

Required documents to be submitted for project management

Project management plan (1 document as a group is needed.)

(although not required, you can use the "project\_schedule\_Template2.xlsx" & "work\_breakdown\_structure\_template2.xlsx".)

your group can turn in either "project schedule" or "work breakdown" or both.

Meeting minutes (1 document as a group is needed.)

(although not required, you can use the "Group Project Meeting Minutes Template.docx".)

(Submission instruction)

(if doing extra credit) Email the 1st "Project management plan" (either project schedule document or work breakdown structure) to syun15@jhmi.edu by 11:59pm, 3/19/18.

Post part1 rough draft of the group project 2 to Discussion board by 11:59pm, 3/19/18

Submit (all the extra credit documents) along with actual group project files at the Blackboard by the deadline (4/7/18).