1. Create a diagram of your region with a detailed figure legend. It should include all genes in your region. You can use UCSC, Ensembl or IGV to create the image (or anything else that works)  
  
*Note: I have attached a figure describing a gene cluster in a bacterium. That's not what you are describing, but I wanted to give you an example of a \*somewhat\* detailed figure legend. Notice how details such as arrows in the picture are explained. Your figure legend should be at least as detailed. Quite often, readers look at the figures before deciding whether or not to read the full paper, so tell the reader exactly what he/she is looking at.*

(Q: Do we need to draw the figure or just screen capture is fine?

A: We recommend drawing the figure as it will demonstrate solid understanding. However, we will accept screen capture if your group choose to do so.)

2. Create a diagram (any method) of all of the transcripts in your gene. You should be zoomed in on the region of your gene for this one. Make sure they are well-labeled either in the diagram and described in the figure legend. You should be able to see the exon patterns, the coding regions and whether or not a protein is produced for each known transcript.  
  
3. In about 500-1000 words, what would you tell someone who is unfamiliar with this gene about this gene (specifically gene product)?  
  
4. Use Galaxy or Biomart to find out something about the noncoding or other regions in your chromosomal region. You can choose what you select to look for in the non-coding regions. You pretty much have anything in the Table Browser/Ensembl from which to choose. Submit the results from Galaxy (or Biomart) and a short explanation what it is you searched for (and what that means).  
  
By about 02/15/18, post your part 1 solutions (even if it's in draft form) in the Discussion Forum for the Group Project. You will be able to see how the other groups approached the problem and you can borrow ideas from each other.  
  
**Part Two - Novel SNPs in your gene**  
I recommend not even starting Part Two until you are at least well on your way to completing Part One. Each group will get a 2016 research paper describing novel SNPs (or variants) in your gene. The papers are below:

|  |  |  |
| --- | --- | --- |
| **Group (Team)** | **PubMed ID** | **Variant** |
| **1** | **27612597** | **G93E** |
| **2** | **27586135** | **S1091\*** |
| **3** | **27498682** | **R630W** |
| **4** | **27485918** | **N355H** |
| **5** | **27686364** | **C694Y** |
|  |  |  |

1, Use the figure with the transcripts from your gene that you developed in Part One. In a separate diagram, label your SNP location on each transcript. Your SNP may not be present on each transcript. Your figure legend should include how each transcript that has the SNP is affected (e.g. change to protein at location X, 5' UTR, noncoding transcript, etc.).  
  
2. Compare the phenotype of your variation to other known variations that affect the phenotype of this gene. About 500-700 words should be about right. OMIM could be a good source here.  
  
3. Use Galaxy, Biomart and IGV to find SNPs and CNVs in your gene. Report your results and in about 500-700 words, compare and contrast what types of results you were able to get from each.  
  
  
  
**Group Memberships**  
You can divide the work any way, but I recommend a lot of overlap. It's a bit of work, but you have time and you can collaborate. You can use your group wikis or Google Docs to share work.  
  
At the end of the project, everyone will get a short survey asking you to assess your contribution and the contributions of your group mates.