Project Outline

This project assignment constitutes 50% of your final grade.
Project reports will be individually graded, so each of you must submit a separate report. You have already been assigned your group, please make sure you meet as soon as possible to discuss complementary skills and to find common research interests.

The finished project and reports are due on Dec. 3rd.
The main purpose of this project is for you to work with colleagues from different disciplinary approaches to conduct a bioinformatics analysis and present the results. The project assignments are in two parts, each part has its guidelines that you must follow. In summary, the parts will cover:

1. Report: Background, Data, Preliminary analysis – points go towards ability to formulate a bioinformatics question, know where to go for the data, and how to download and store it. If you have already started analysis, that's fine too.

2. Presentation: Analysis Results, interpretation, presentation – points go towards ability to analyze the data, and understand the results. This part also includes the synthesis, which is the presentation at the end where it all comes together. Points go towards ability to interpret the data and formulate next steps.

Part 1. Report: background, data, prelim analysis

1. Background. Start with who your colleagues are, the type of research they are involved with and whether their disciplinary approach is at a wet lab or computational. Next, outline the purpose of your project in one paragraph. Why is it important and what are you trying to achieve. Try to use a data driven approach, rather than being too specific in your hypotheses.

2. Data. Download data such as RNAseq or genomic data from one of publicly available databases. You can use any data mining method available to you, whether learnt in the class or already known from you or your colleagues' (graduate) work, but you must fully describe the data and download method. You are not limited to the human genome, but please do not work on an organism that has not been fully annotated (Drosophila and mice are OK, try to avoid strange sea creatures).

3. Analysis (make sure you write about your methods in detail).

First you will conduct basic analysis such as mapping to a genome and/or other types of sequence alignments.

Then you will take your analysis one step further, in any direction that is relevant to your biological question. You can, for example, compare the miRNA and gene expression data sets of a particular experiment or disease condition, or other non-coding RNA analysis on your gene set. You can look at epigenetic patterns across the whole genomic region where your gene of interest lies, and make predictions regarding time and space of transcription based on your findings.

It is fairly open to what you do in this part, as long as it is within computational analysis, is not elementary, and helps answer the biological question. You may have noticed, your graded homework assignments are pretty specific, here you are more free to create your own project, as long as you follow the guidelines and use the tools you are learning about in class.

Report Guidelines

- Please submit reports of 2-3 pages long, 1.5 or double spaced, written in Times or Calibri, font 12, not counting tables and figures. Neat presentation is required throughout. Captions must be used for any and all tables and figures.
- Any references must be on a separate, additional, page.
- Any plagiarism, from articles, websites, or each other, will result in a 0 grade.
- Late submissions will not be accepted.

Part 2. Presentation: results, interpretation and presentation
Your written reports (pdf or docx) and presentation (pdf or pptx) files are due by email before class on December 3rd.

You are expected to present for 15 minutes on Tuesday December 3rd with an additional 5 minutes for questions. The order of presenting teams will be chosen at random, so arrive on time and arrive prepared.

Start with a brief introduction about the question you have set out to answer. Then describe your methods, your results, and then interpret the results not only in the context of the research area you worked in, but also in the context of the broader research community. The final slide should be a synthesis about what you would do next in this project.

The methods have to be computationally sound. The interpretation has to make sense. The synthesis can include computational work using existing tools, it can include a proposal for an algorithm or software that you would like to develop that would help this sort of analysis, and, if it is logical to go back to the lab with this project, then that is also a valid synthesis.

The presentation must include the following (number of slides provided as a suggested guideline):

- a title (1 slide)
- an introduction to the topic (1 slide)
- the question you have set out to answer (1 slide)
- methods (2-3 slides)
- results (3-4 slides)
- interpretation of the results, ie how this has answered the question you posed in #3 (1-2 slides)
- synthesis, i.e. how this is informing what you do next (1 slide)