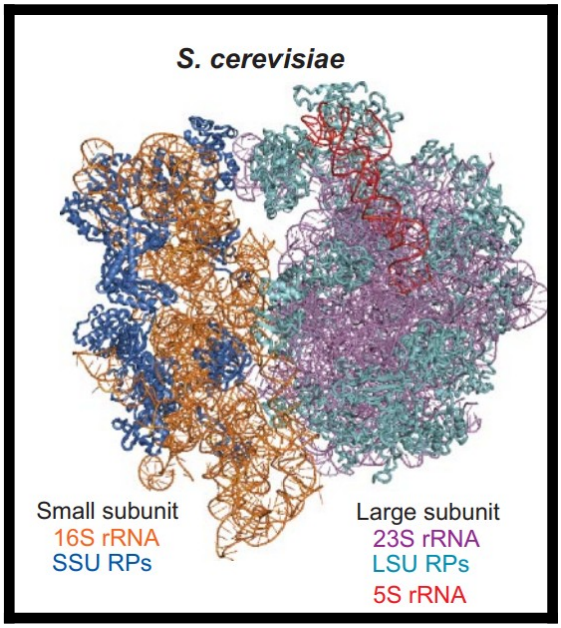
Lab 8: R packages

Biol135 Introduction to Bioinformatics Programming




***S. Cerevisiae* Ribosome.**

For this lab we will incorporate the use of some R packages into our Dinucleotide Transition Matrix

Part 0: In Rstudio create a new script called [your last name]transitionPlot.R

Part1: Using your counter from Lab 7 quantify the dinucleotides from a real sequence taken from Saccharomyces Genome database. We will use the genomic DNA Sequence in Fasta (.fsa) format:

https://www.yeastgenome.org/locus/S000006487

Download the sequence and store it in your working directory.

Part 1:  Use the read.fasta command in seqinr to load your sequence from the file.

Part 2:

Install and load the ggplot2  and seqinr packages. Using ggplot2  You will need to structure your dinucleotide frequencies in a way that they can be plotted using the geom_tile() command in ggplot2. Make sure your plot has a title, axis titles, and that the color in the heatmap correspond to the frequency of the dinucleotide observed in your sequence.

Part 3: Add comments to your code and upload to the MyCourses Dropbox.