After I started working on the yeast project, various small details needed to be figured out before I could go further with the project. From data collection to how to analyze the data, these are all things that I have been working on. Originally, what I had was a list of all SNPs in predicted TFBSs. There is also information regarding genes that exhibit differences in gene expression between two yeast strains (RM and BY) which are very likely due to cis regulatory SNPs. Even though I had already done part of the analysis, I had to restart it gain because I realized that I should be using another data. At the moment I have calculated shape changes for all SNPs, and the next step is to compare the different subgroups (change in shape, non change in shape, cis, non cis). As I said before, even if changes in shape are not statistically different between the two groups, we could still try to find potential candidates in which would be interesting to study TF binding that is influenced by different DNA shapes. I also pulled out from the Transfac a list of human SNPs associated with some sort of disease/phenotype that are in TFBSs. With that, I can investigate shape changes and depending on the TF involved, elucidate the mechanism of changes in DNA shape readout due to SNPs for some phenotypes. This week I hope to finish analyzing the yeast data and start with the transfac one.