Carlos Gonzalez

10/24/2016

2) the SELEX-seq and PBM of in vitro models allow us to use a linear regression line to compare the similarities if a certain of molecule we are looking for. One of the advantages that this method has is that we can use this model for a more specific approach. Another advantage I that we can encode Dna sequences using binary codes and from using binary code we can then find the feature vector of each base which will reflect the shape of the dna including propeller twists, helical twists etc. one disadvantage is that the model is limited to x- amount of samples The in vivo experiment CHIP-seq relies more of the classification of the model whether it is a good model or a bad model. This model uses the are underneath the curve of an ROC curve to justify the model presented. One advantage of this model is that based o the graph you can tell whether the model is correct or not. One disadvantage is that this model is limited to only analyzing what is already processed by the SELEX-seq.

5) b. I learned based on the results that although the point were a bit off the regression line, they were all similar based on the shape of the molecules. This leads to the point that through amplification of a certain molecule through the SELEx-seq, the error is insignificant.

8) although the regression line was great, the ROC curve was horrible. Maybe because there was some coding error however, it showed that my model was not great but in fact bad with an AUC of less than .5