**Methods**

The first part of the challenge requires predicting the log fold changes of the reference and alternate allele sequences a well the regulatory hit probability.

For the regulatory hit prediction, we used k-mers of length 4 as features and used an SVM for binary classification. We then performed 5 fold cross validation and computed the ROC which is the shown in Figure 1. An average AUC of 0.63 was achieved.

For predicting the log fold change, we used k-mers of length 4 as features and used SVM regression. Again we performed 5 fold cross valuation and recorded values, such as, correlation and rmse. The results are shown in Figure 2 which shows and rmse of 0.36 and a correlation of 0.33.

****

**Fig 1: ROC for regulatory hit prediction**

****

Actual logfold change

Predicted logfold change

**Fig 2: Predicted vs actual log fold change for the reference sequences**