

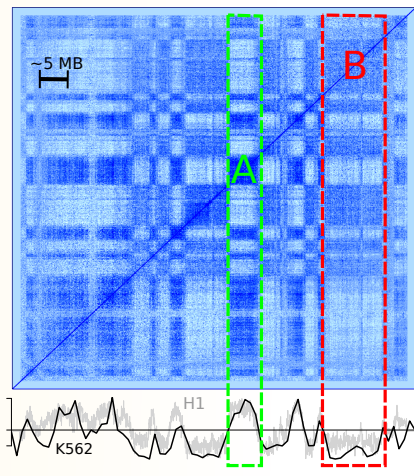
**Eigenvectors or interaction matrices** from:  
 K562, GM06690: Lieberman-Aiden *et al.* (2009)  
 H1, IMR90: Dixon *et al.* (2012)  
 GM12878: Kalhor *et al.* (2011)

Calculate principal component eigenvectors from normalised interaction matrices. Select chromosomes with meaningful eigs across each cell type.

ChIP-Seq data (MACS2 control-normalised signal) generated by the **ENCODE** consortium including 11 histone modifications and 22 ChIP-seq DNA-binding protein datasets

Find average values for each ENCODE dataset across matched 1MB bins for K562, H1 and GM12878.

H1 hESC chromosome 18 — 40 kb interaction matrix

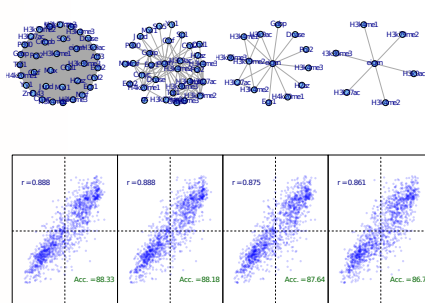


**Graphical LASSO** regularisation with varying lambda.

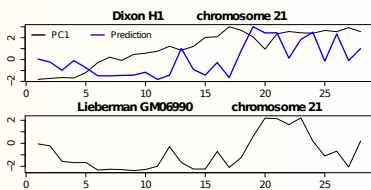
**Random Forest** model  
 eig output ↔ Feature set  
 10-fold cross validation

Evaluate performance

L1 penalty: 0 L1 penalty: 0.01 L1 penalty: 0.1 L1 penalty: 0.2



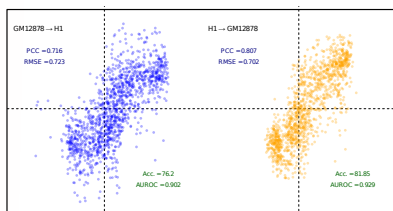
Predict eigs for excluded chromosomes



Compare variable importance

Cross-apply models to different cell types

GM12878 → H1 model features  
 H1 → GM12878 model features



Modelling

Evaluation