

Summary reports of non-classic function detection of : test3

January 6, 2019

Contents

1	Data description	2
2	ElasticNet co-factor selection	3
3	potential co-factors corresponded to non-classic function	4
3.1	summary of co-factors	4
3.2	Boxplot of HM on non-classic and classic sites	5
4	Output list	6

1 Data description

Table 1 mainly describes the input files, parameters and options.

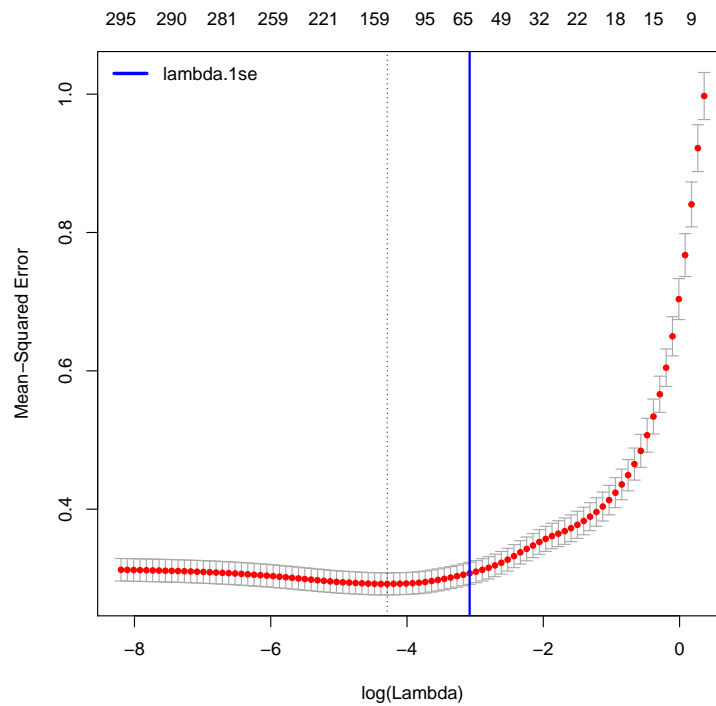
Table 1: parameter description

parameter	value
output name	test3
HMRpeak(peak filename)	mESC_GSM1562337_CBX7.bed
HM signal(bw filename)	mESC_GSM1399500_H3K27me3.bw
#coTF candidates	143
options	value
extend size	1000bp
Alpha (Elastic net)	0.5
Pvalue cutoff	0.001
topN cofactors	all

2 ElasticNet co-factor selection

In this step we use a feature selection (elastic-net. Zou, H. and Hastie T. (2005)) to select potential co-factors which corresponded to the non-classic function. Below shows the cross-validation curve for the decision of lambda in elastic-net.

Figure 1: cross-validation curve for lambda decision



3 potential co-factors corresponded to non-classic function

In summary, 60 factors were predicted to potentially act as a co-factor of the non-classic function. The top13 co-factors were listed.

3.1 summary of co-factors

The empirical p-value, R-square (ordered) and the number of non-classic (NC) sites for each potential co-factors were listed below. The empirical p-value was calculated based on the comparison of foreground (observed) R-square and background R-square (distribution of random R-square generated from the 1,000 permutations of co-binding events) for each potential co-factor. The non-classic sites were defined by lower HM signal (using Otus' method) and co-binding events of each potential co-factor.

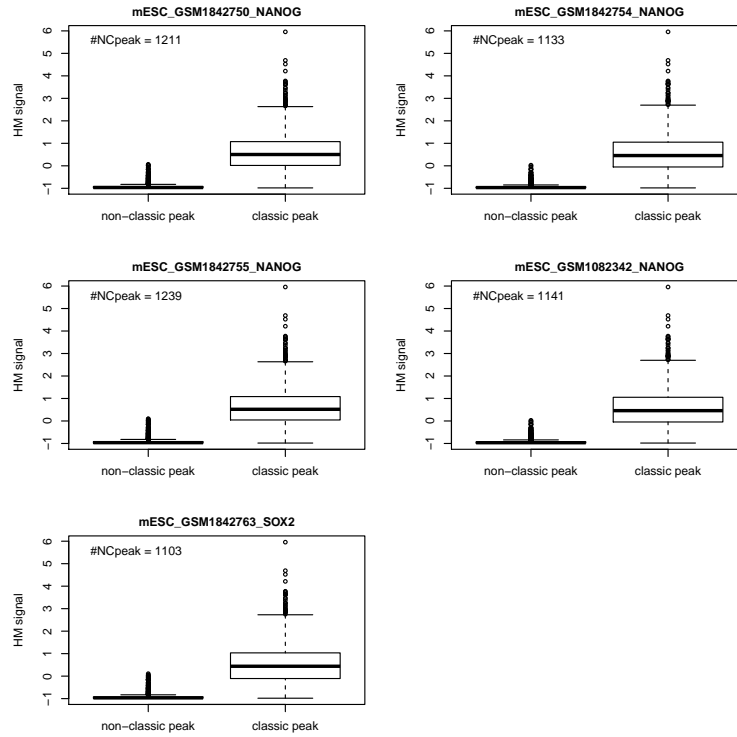
Table 2: cofactor summary

co-factor	p-value	R-square	#NCsites
mESC_GSM1842750_NANOG	0.001	0.489	1211
mESC_GSM1842754_NANOG	0.001	0.478	1133
mESC_GSM1842755_NANOG	0.001	0.469	1239
mESC_GSM1082342_NANOG	0.001	0.466	1141
mESC_GSM1842763_SOX2	0.001	0.375	1103
mESC_GSM623989_PRDM14	0.001	0.349	1001
mESC_GSM1355155_POU5F1	0.001	0.311	1122
mESC_GSM1019073_EP300	0.001	0.245	752
mESC_GSM1208218_KLF5	0.001	0.162	761
mESC_GSM1047933_POU5F1	0.001	0.15	629
mESC_GSM1562331_POU5F1	0.001	0.116	464
mESC_GSM1406445_TRIM28	0.001	0.116	568
mESC_GSM560348_MED1	0.001	0.111	546

3.2 Boxplot of HM on non-classic and classic sites

Boxplot was generated to compare the difference of the histone mark (HM) signal on either non-classic(NCpeak) or classic sites. The non-classic sites were defined by lower HM signal (using Otus' method) and co-binding events of each potential co-factor. The boxplot corresponded to top5 co-factors were displayed.

Figure 2: boxplot cofactor HMsignal



4 Output list

All output files were described in the following table

Table 3: output list

description	filename
cobinding matrix on HMR peaks	tmpResults/test3_peakov.bed
histone mark signal on HMR peaks	tmpResults/test3_HMsig.bed
summary table of non-classic function	summary/test3_NCsummary.txt
summary report (this doc)	summary/test3_summary.pdf