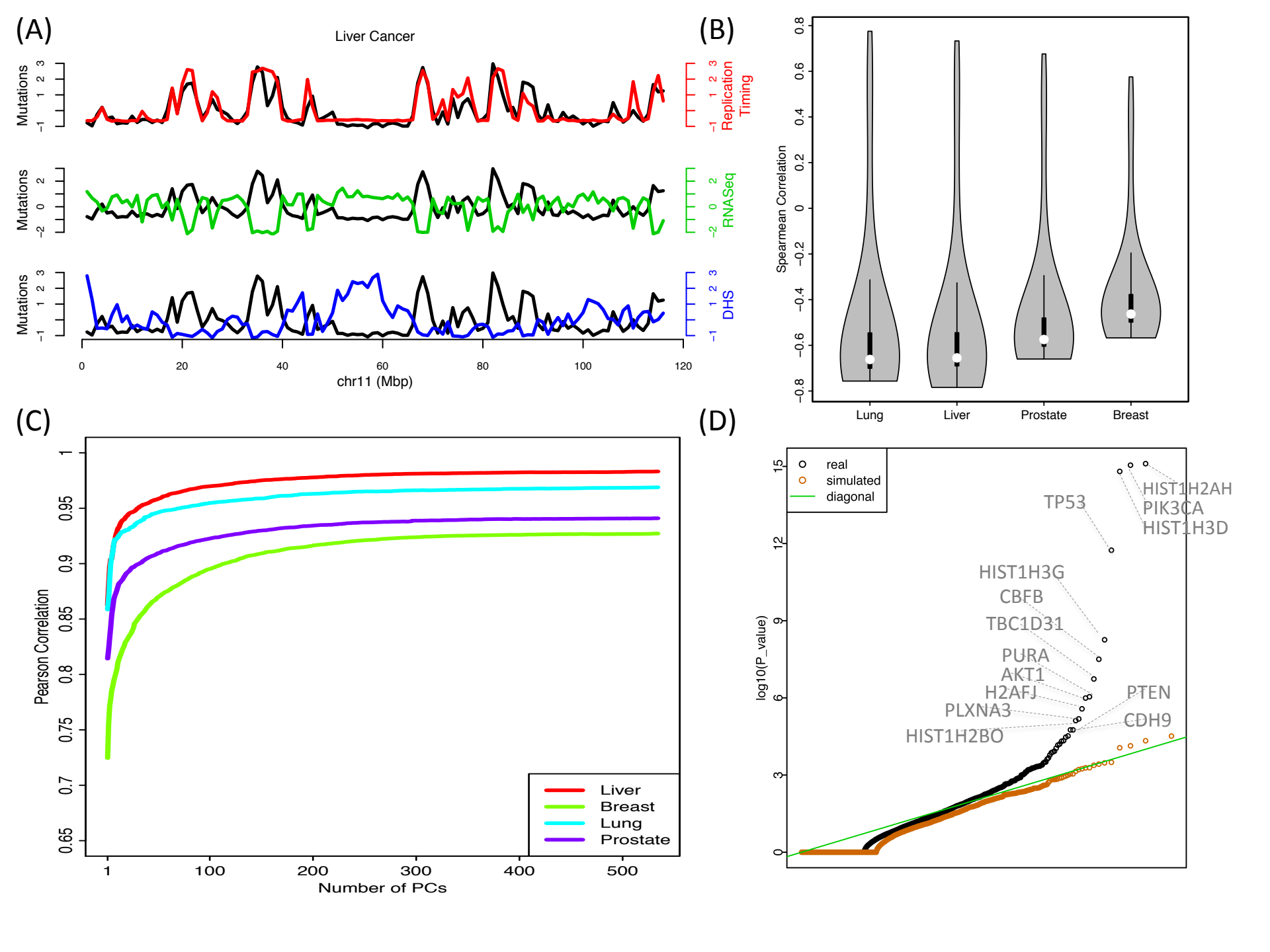
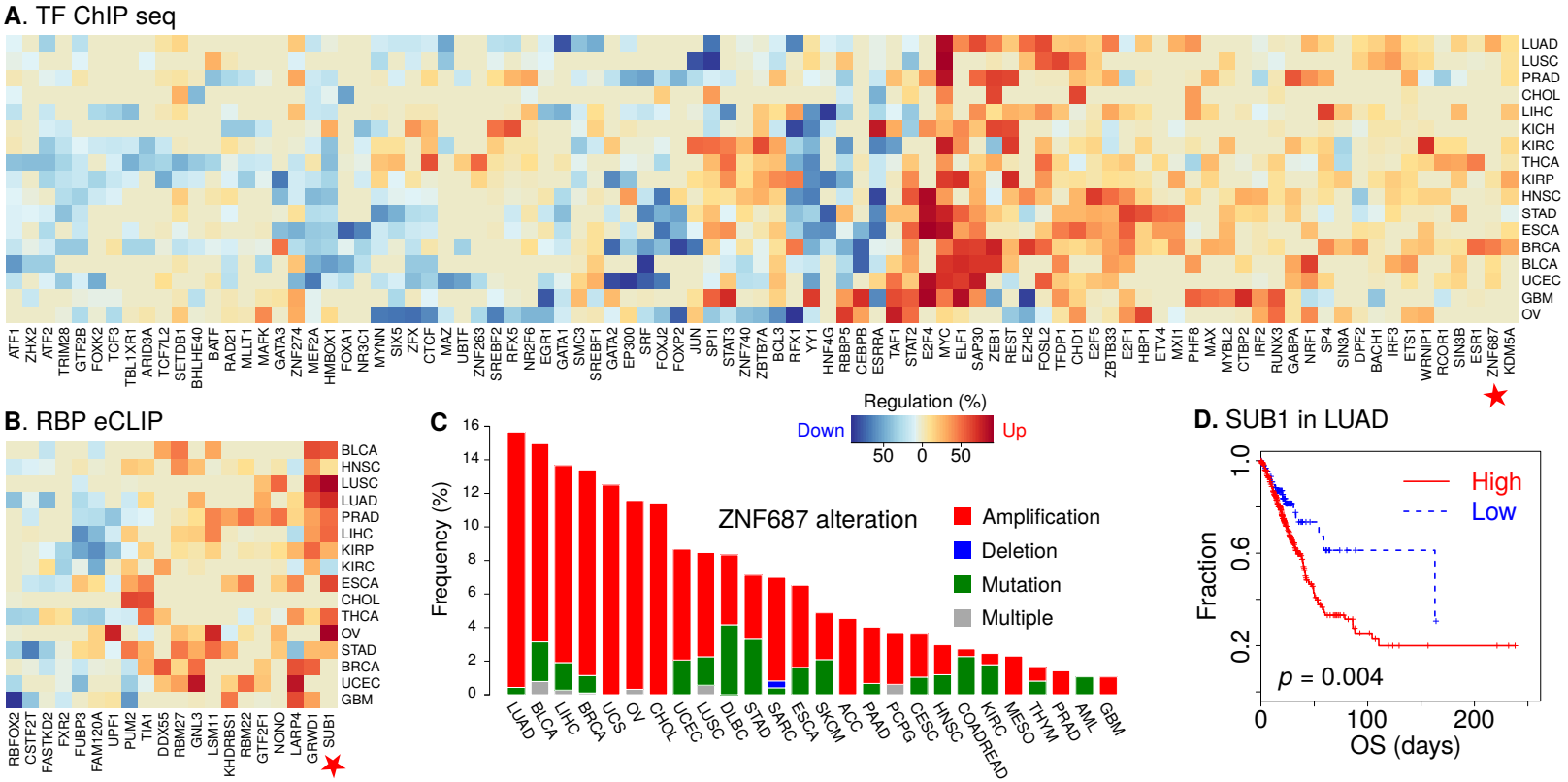
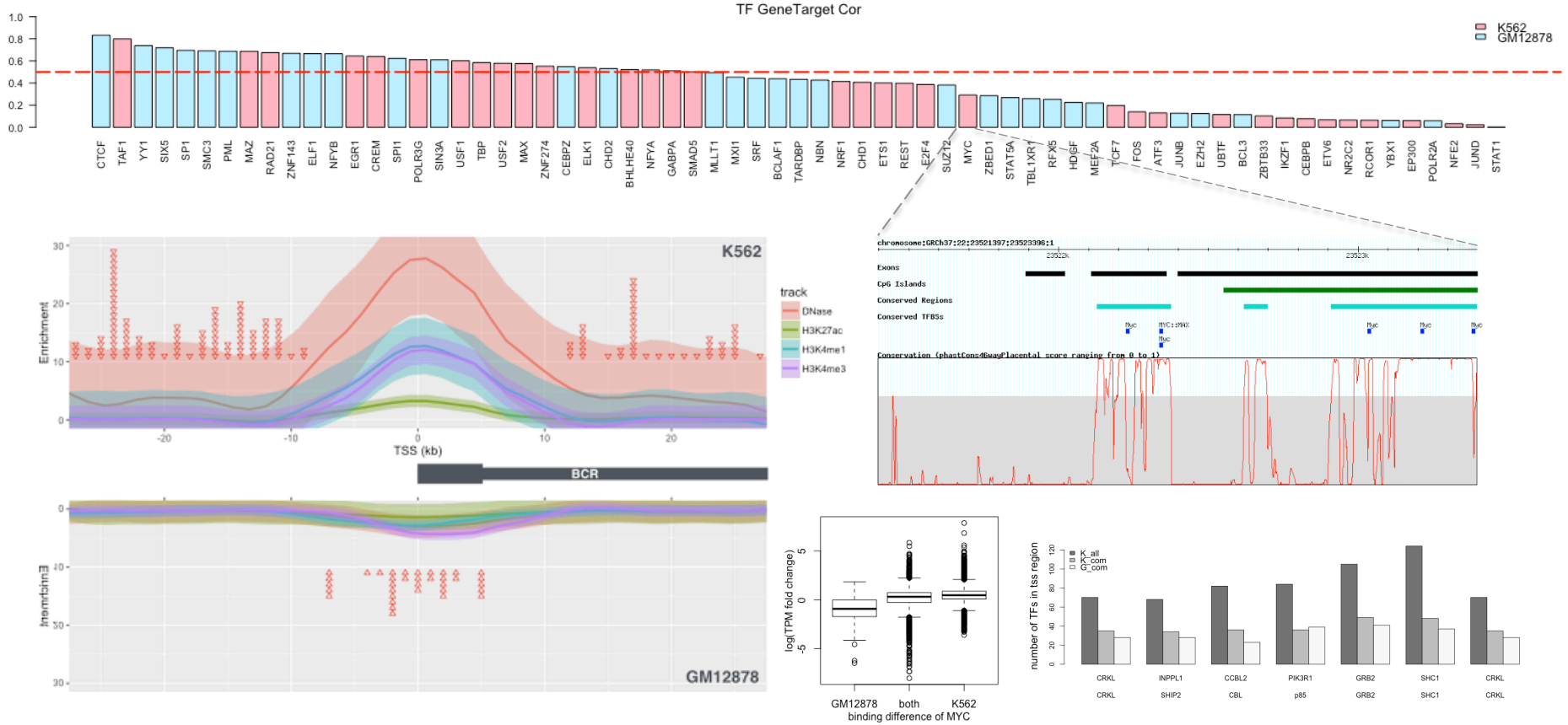
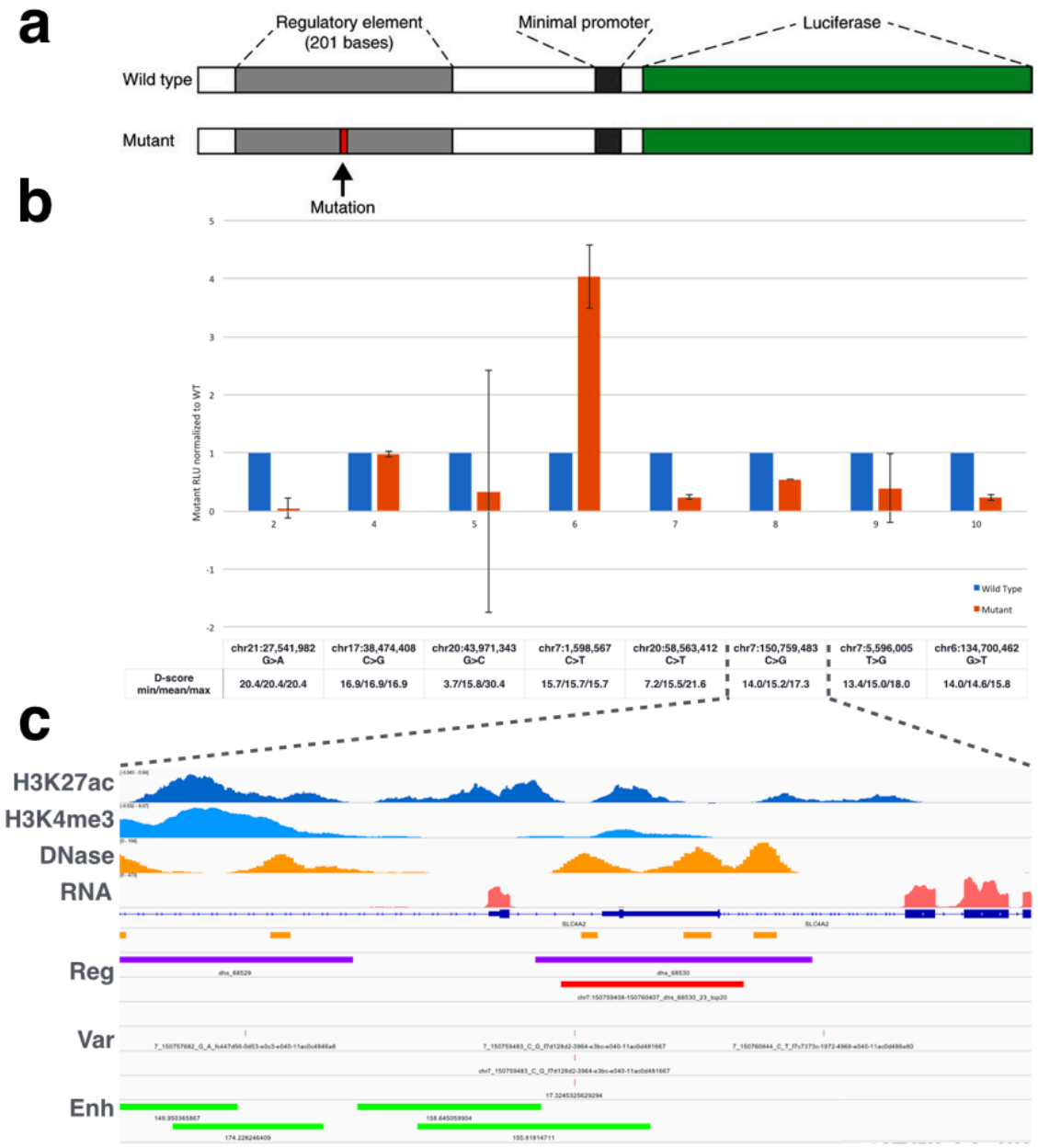
Encode Cancer Figure Slides


(A)
Liver Cancer
Mutations
Replication Timing
RNASeq
DHS
chr11 (Mbp)
(B)
Spearmean Correlation
Lung
Liver
Prostate
Breast
(C)
Pearson Correlation
Number of PCs
Liver
Breast
Lung
Prostate
(D)
real
simulated
diagonal
log10(P_value)
TP53
HIST1H2AH
PIK3CA
HIST1H3D
HIST1H3G
CBFB
TBC1D31
PURA
AKT1
H2AFJ
PLXNA3
HIST1H2BO
PTEN
CDH9



Inference of regulators driving tumor specific expression patterns. Based on ENCODE ChIPSeq and eCLIP data, we applied RABIT framework to identify transcription factors (TF) and RNA binding proteins (RBP), whose target genes are differentially regulated in diverse TCGA cancer types. (A) Hierarchically clustered heatmap was used to show the percentage of patients with TF target significantly up-regulated (red) or down-regulated (blue). Only TFs with targets up-regulated in over 30% patients in at least two cancer types are included. (B) The percentage of patients with RBP target differentially regulated was shown by heatmap as part A. (C) The cBioportal alteration plots was shown for *ZNF687*. (D) All TCGA lung adenocarcinoma patients are divided to two groups according to the *SUB1* activity predicted by RABIT. The overall survival was shown in each group by KM plot. The association between RABIT regulatory activity and overall survival was tested CoxPH regression.


TF GeneTarget Cor
K562
GM12878
CTCF TAF1 YY1 SIX5 SP1 SMC3 PML MAZ RAD21 ZNF143 ELF1 NFYB EGR1 CREM SPI1 POLR3G SIN3A USF1 TBP USF2 MAX ZNF274 CEBPZ ELK1 CHD2 BHLHE40 NFYA GABPA SMAD5 MLLT1 MXI1 SRF BCLAF1 TARDBP NBN NRF1 CHD1 ETS1 REST E2F4 SUZ12 MYC ZBED1 STAT5A TBL1XR1 RFX5 HDGF MEF2A TCF7 FOS ATF3 JUNB EZH2 UBTF BCL3 ZBTB33 IKZF1 CEBPB ETV6 NR2C2 RCOR1 YBX1 EP300 POLR2A NFE2 JUND STAT1
K562
Enrichment
TSS (kb)
track
DNase
H3K27ac
H3K4me1
H3K4me3
BCR
GM12878
chromosome:GRCh37:22:23521397:23523396:1
Exons
CpG Islands
Conserved Regions
Conserved TFBSs
Conservation (phastCons46wayPlacental score ranging from 0 to 1)
log(TPM fold change)
GM12878
both
K562
binding difference of MYC
number of TFs in tss region
K_all
K_com
G_com
CRKL
INPPL1
CCBL2
PIK3R1
GRB2
SHC1
CRKL
CRKL
SHIP2
CBL
p85
GRB2
SHC1
CRKL


a
Regulatory element
(201 bases)
Minimal promoter
Luciferase
Wild type
Mutant
Mutation
b
Mutant RLU normalized to WT
Wild Type
Mutant
chr21:27,541,982 G>A
chr17:38,474,408 C>G
chr20:43,971,343 G>C
chr7:1,598,567 C>T
chr20:58,563,412 C>T
chr7:150,759,483 C>G
chr7:5,596,005 T>G
chr6:134,700,462 G>T
D-score min/mean/max
20.4/20.4/20.4
16.9/16.9/16.9
3.7/15.8/30.4
15.7/15.7/15.7
7.2/15.5/21.6
14.0/15.2/17.3
13.4/15.0/18.0
14.0/14.6/15.8
c
H3K27ac
H3K4me3
DNase
RNA
Reg
Var
Enh