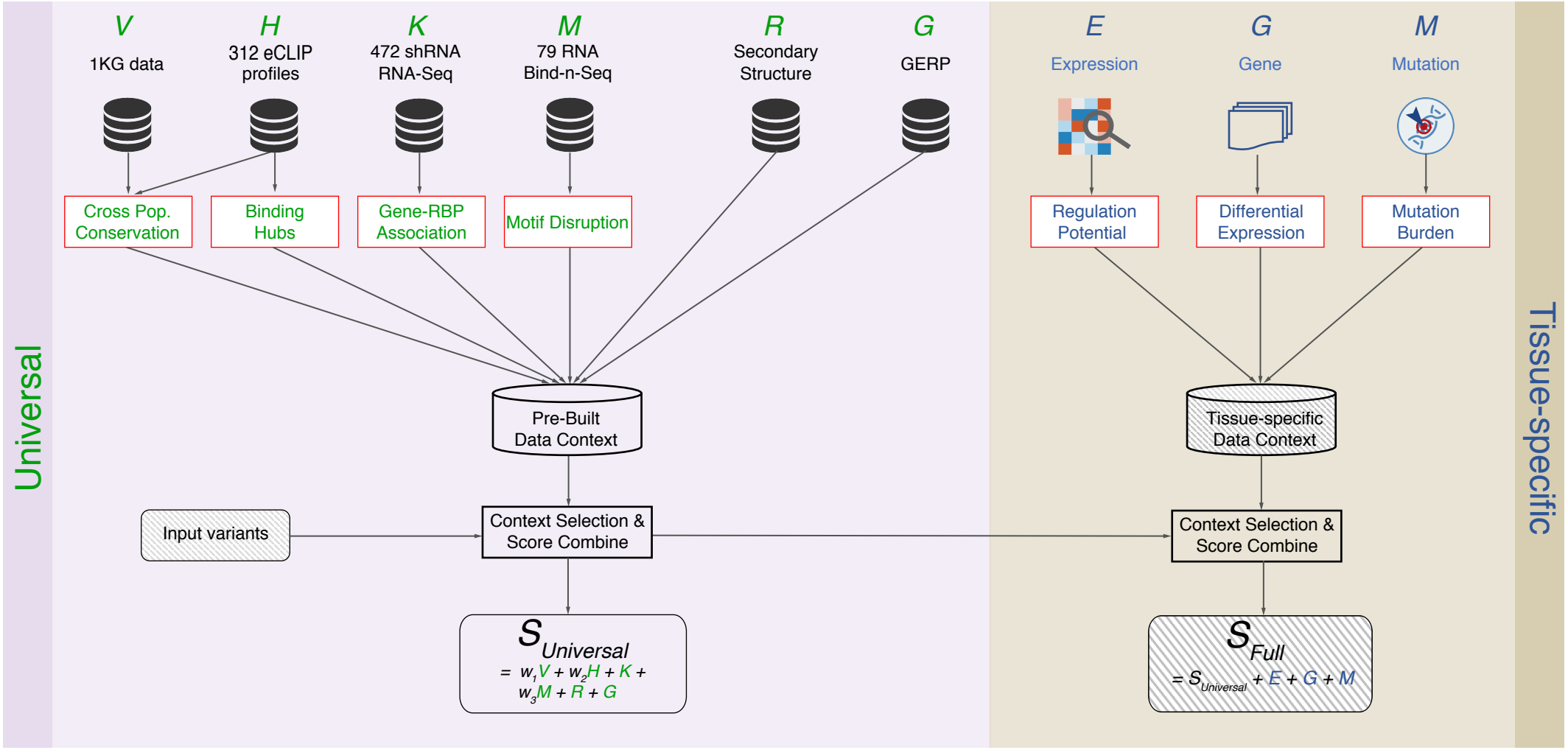
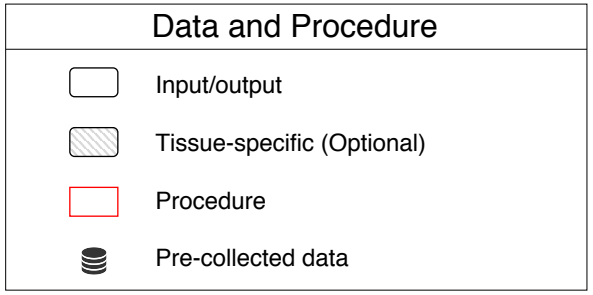
Figure 1


Universal
V
1KG data
H
312 eCLIP profiles
K
472 shRNA RNA-Seq
M
79 RNA Bind-n-Seq
R
Secondary Structure
G
GERP
Cross Pop. Conservation
Binding Hubs
Gene-RBP Association
Motif Disruption
Pre-Built Data Context
Input variants
Context Selection & Score Combine
$S_{Universal} = w_1V + w_2H + K + w_3M + R + G$
E
Expression
G
Gene
M
Mutation
Regulation Potential
Differential Expression
Mutation Burden
Tissue-specific Data Context
Context Selection & Score Combine
$S_{Full} = S_{Universal} + E + G + M$
Tissue-specific


Data and Procedure
Input/output
Tissue-specific (Optional)
Procedure
Pre-collected data

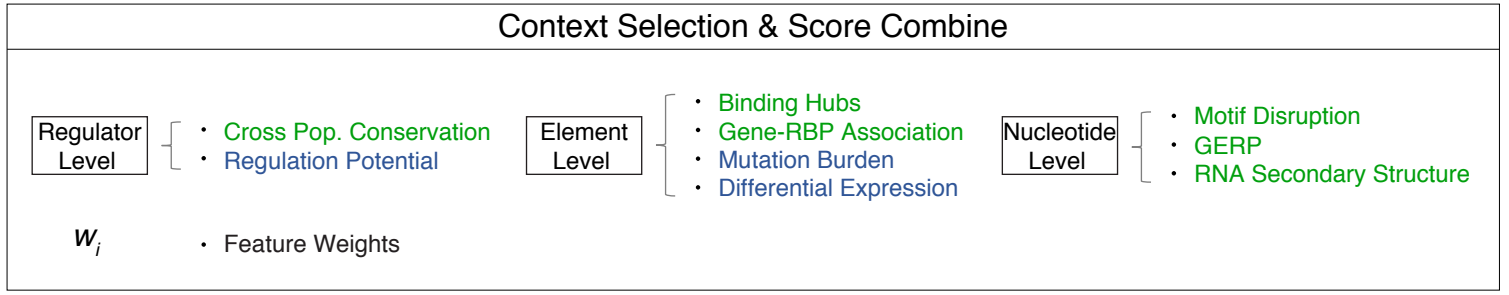

Context Selection & Score Combine
Regulator Level
Cross Pop. Conservation
Regulation Potential
Element Level
Binding Hubs
Gene-RBP Association
Mutation Burden
Differential Expression
Nucleotide Level
Motif Disruption
GERP
RNA Secondary Structure
w_i
Feature Weights

Figure 2

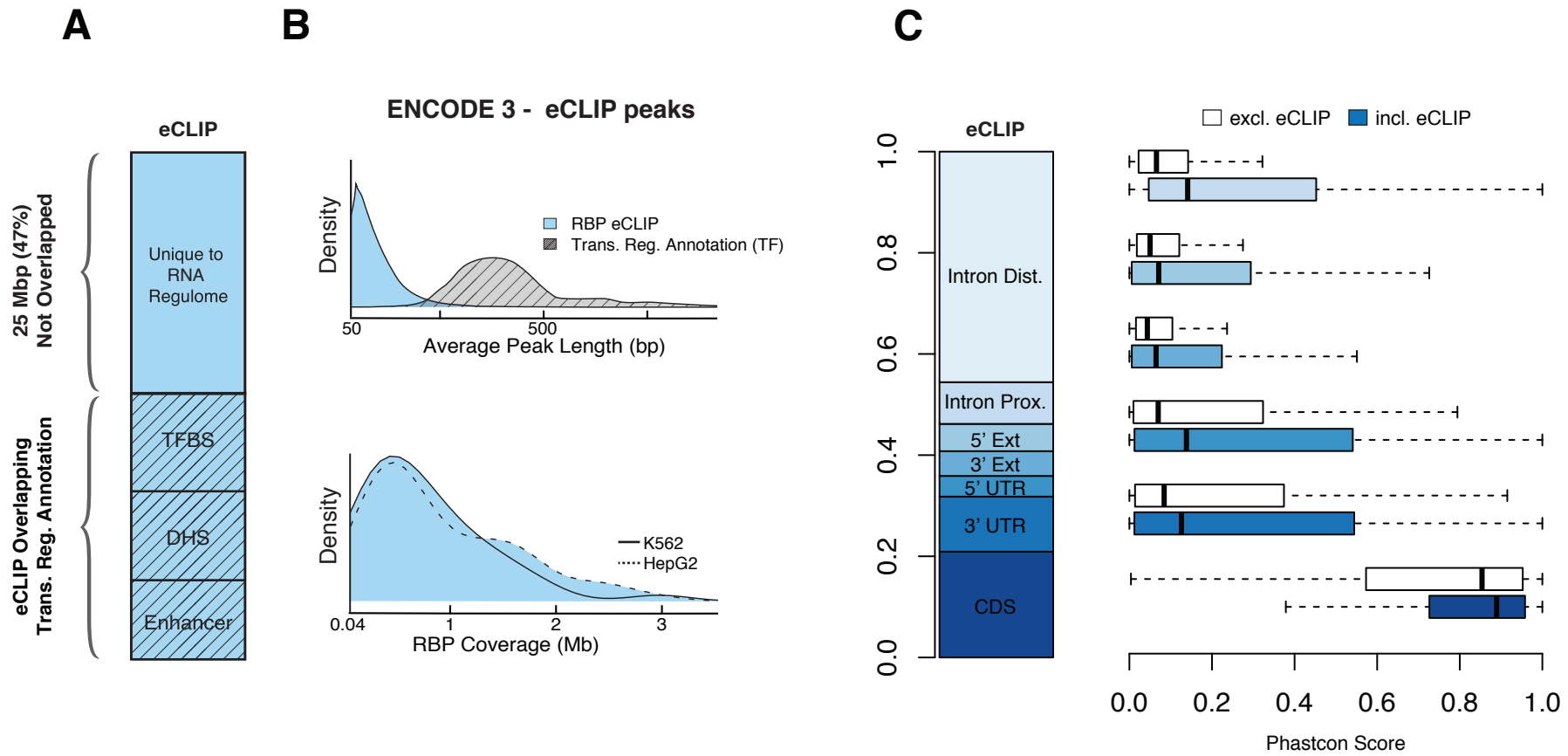

A
B
C
eCLIP
25 Mbp (47%)
Not Overlapped
eCLIP Overlapping
Trans. Reg. Annotation
Unique to
RNA
Regulome
TFBS
DHS
Enhancer
ENCODE 3 - eCLIP peaks
Density
RBP eCLIP
Trans. Reg. Annotation (TF)
50
500
Average Peak Length (bp)
Density
K562
HepG2
0.04
1
2
3
RBP Coverage (Mb)
eCLIP
Intron Dist.
Intron Prox.
5' Ext
3' Ext
5' UTR
3' UTR
CDS
0.0
0.2
0.4
0.6
0.8
1.0
excl. eCLIP
incl. eCLIP
0.0
0.2
0.4
0.6
0.8
1.0
Phastcon Score

Figure 3

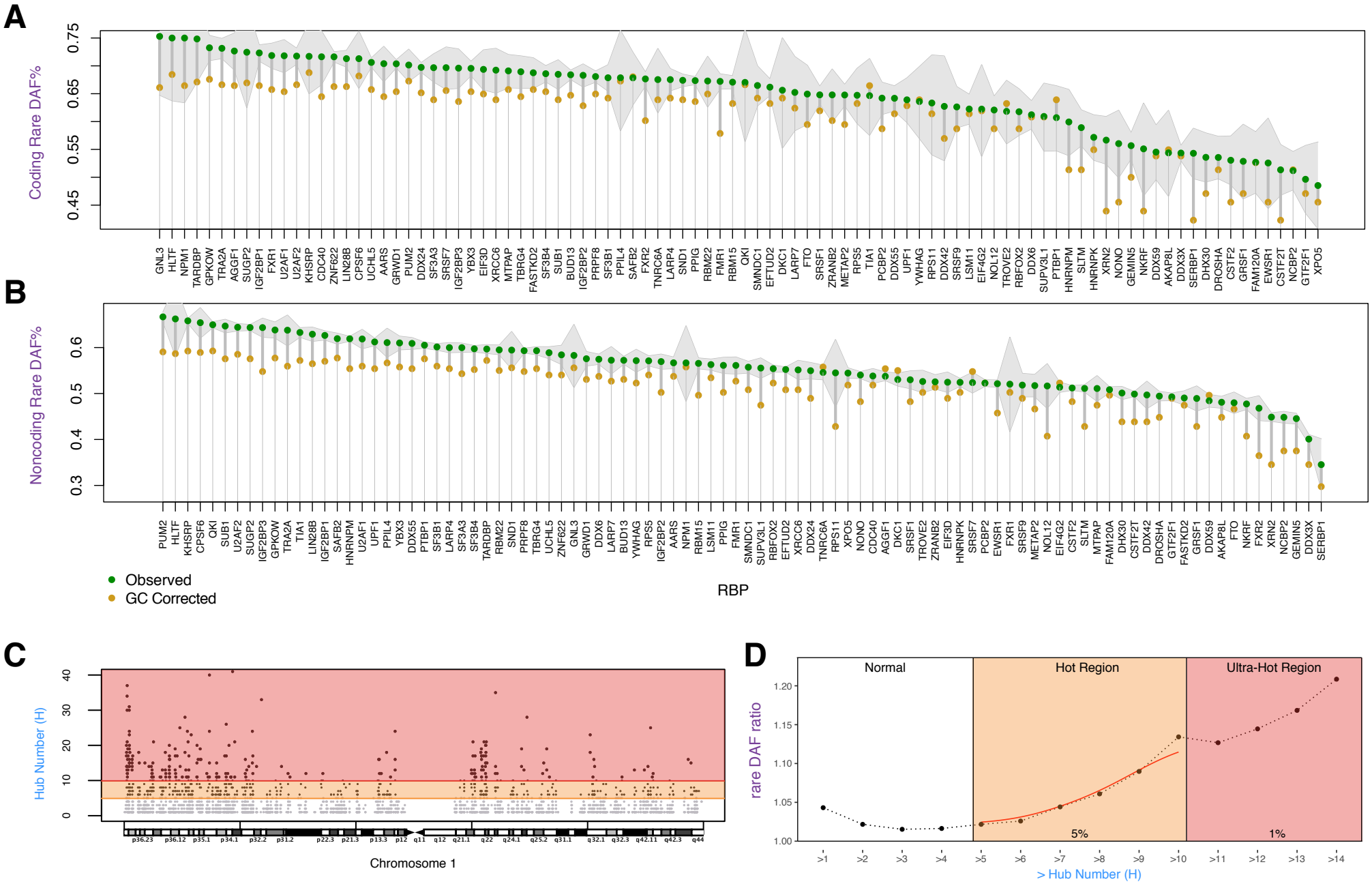

A
Coding Rare DAF%
0.45
0.55
0.65
0.75
GNL3 HLTF NPM1 TARDBP GPKOW TRA2A AGGF1 SUGP2 IGF2BP1 FXR1 U2AF1 U2AF2 KHSRP CDC40 ZNF622 LIN28B CPSF6 UCHL5 AARS GRWD1 PUM2 DDX24 SF3A3 SRSF7 IGF2BP3 YBX3 EIF3D XRCC6 MTPAP TBRG4 FASTKD2 SF3B4 SUB1 BUD13 IGF2BP2 PRPF8 SF3B1 PIIL4 SAFB2 FXR2 TNRC6A LARP4 SND1 PPIG RBM22 FMR1 RBM15 QKI SMNDC1 EFTUD2 DKC1 LARP7 FTO SRSF1 ZRANB2 METAP2 RPS5 TIA1 PCBP2 DDX55 UPF1 YWHAG RPS11 DDX42 SRSF9 LSM11 EIF4G2 NOL12 TROVE2 RBFOX2 DDX6 SUPV3L1 PTBP1 HNRNPM SLTM HNRNPK XRN2 NONO GEMIN5 NKRF DDX59 AKAP8L DDX3X SERBP1 DHX30 DROSHA CSTF2 GRSF1 FAM120A EWSR1 CSTF2T NCBP2 GTF2F1 XPO5
B
Noncoding Rare DAF%
0.3
0.4
0.5
0.6
PUM2 HLTF KHSRP CPSF6 QKI SUB1 U2AF2 SUGP2 IGF2BP3 GPKOW TRA2A TIA1 LIN28B IGF2BP1 SAFB2 HNRNPM U2AF1 UPF1 PIIL4 YBX3 DDX55 PTBP1 SF3B1 LARP4 SF3A3 SF3B4 TARDBP RBM22 SND1 PRPF8 TBRG4 UCHL5 ZNF622 GNL3 GRWD1 DDX6 LARP7 BUD13 YWHAG RPS5 IGF2BP2 AARS NPM1 RBM15 LSM11 PPIG FMR1 SMNDC1 SUPV3L1 RBFOX2 EFTUD2 XRCC6 DDX24 TNRC6A RPS11 XPO5 NONO CDC40 AGGF1 DKC1 SRSF1 TROVE2 ZRANB2 EIF3D HNRNPK SRSF7 PCBP2 EWSR1 FXR1 SRSF9 METAP2 NOL12 EIF4G2 CSTF2 SLTM MTPAP FAM120A DHX30 CSTF2T DDX42 DROSHA GTF2F1 FASTKD2 GRSF1 DDX59 AKAP8L FTO NKRF FXR2 XRN2 NCBP2 GEMIN5 DDX3X SERBP1
Observed
GC Corrected
RBP
C
Hub Number (H)
0
10
20
30
40
p36.23 p36.12 p35.1 p34.1 p32.2 p31.2 p22.3 p21.3 p13.3 p12 q11 q12 q21.1 q22 q24.1 q25.2 q31.1 q32.1 q32.3 q42.11 q42.3 q44
Chromosome 1
D
Normal
Hot Region
Ultra-Hot Region
rare DAF ratio
1.00
1.05
1.10
1.15
1.20
5%
1%
>1 >2 >3 >4 >5 >6 >7 >8 >9 >10 >11 >12 >13 >14
> Hub Number (H)

Figure 4

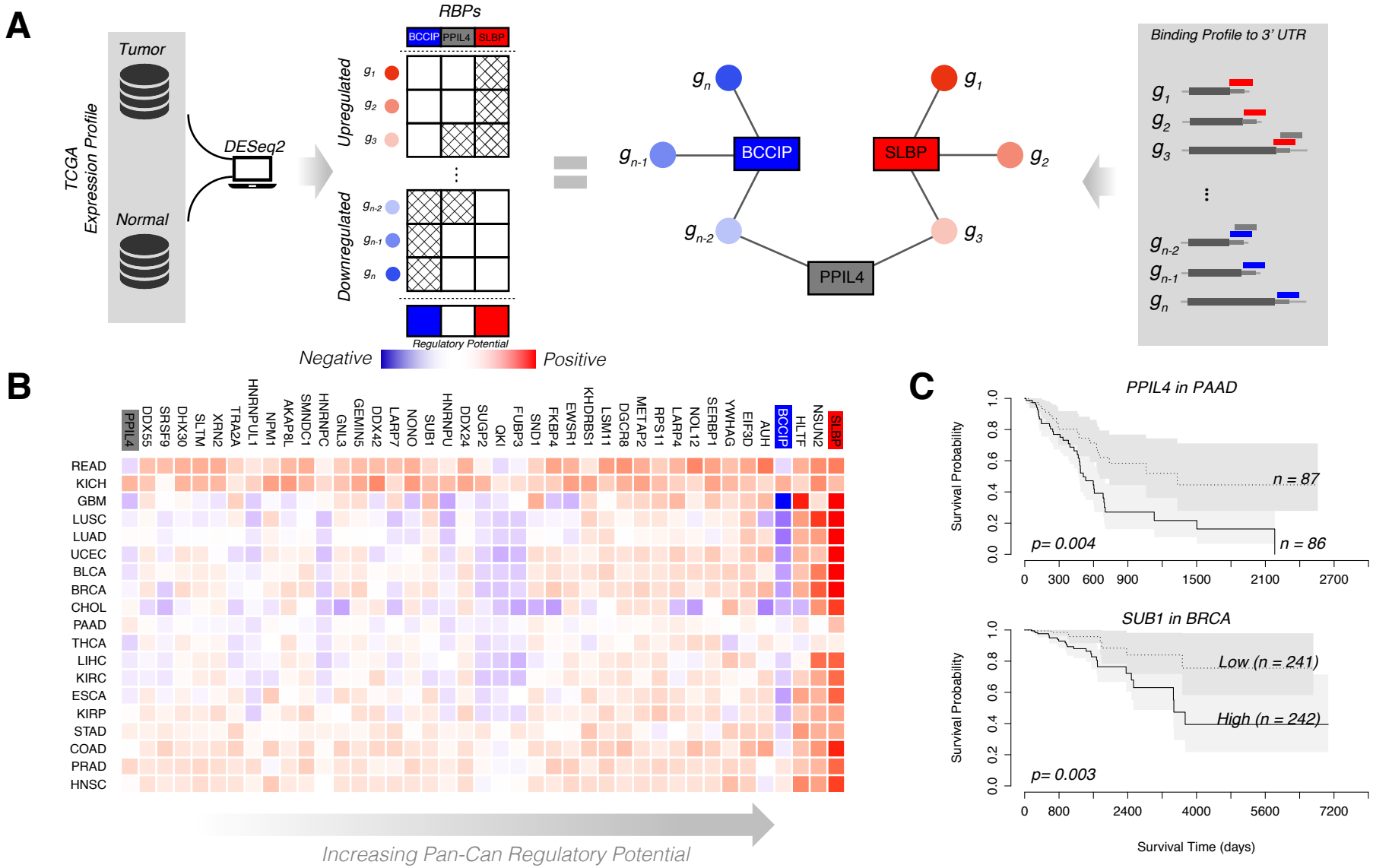

A
Tumor
Normal
TCGA Expression Profile
DESeq2
RBPs
BCCIP
PPIL4
SLBP
Upregulated
Downregulated
Regulatory Potential
Negative
Positive
BCCIP
SLBP
PPIL4
Binding Profile to 3' UTR
B
PPIL4
DDX55
SRSF9
DHX30
SLTM
XRN2
TRA2A
HNRNPUL1
NPM1
AKAP8L
SMNDC1
HNRNPC
GNL3
GEMIN5
DDX42
LARP7
NONO
SUB1
HNRNPU
DDX24
SUGP2
QKI
FUBP3
SND1
FKBP4
EWSR1
KHDRBS1
LSM11
DGCR8
METAP2
RPS11
LARP4
NOL12
SERBP1
YWHAG
EIF3D
AUH
BCCIP
HLTF
NSUN2
SLBP
READ
KICH
GBM
LUSC
LUAD
UCEC
BLCA
BRCA
CHOL
PAAD
THCA
LIHC
KIRC
ESCA
KIRP
STAD
COAD
PRAD
HNSC
Increasing Pan-Can Regulatory Potential
C
PPIL4 in PAAD
Survival Probability
n = 87
n = 86
p= 0.004
SUB1 in BRCA
Low (n = 241)
High (n = 242)
p= 0.003
Survival Time (days)

Figure 5

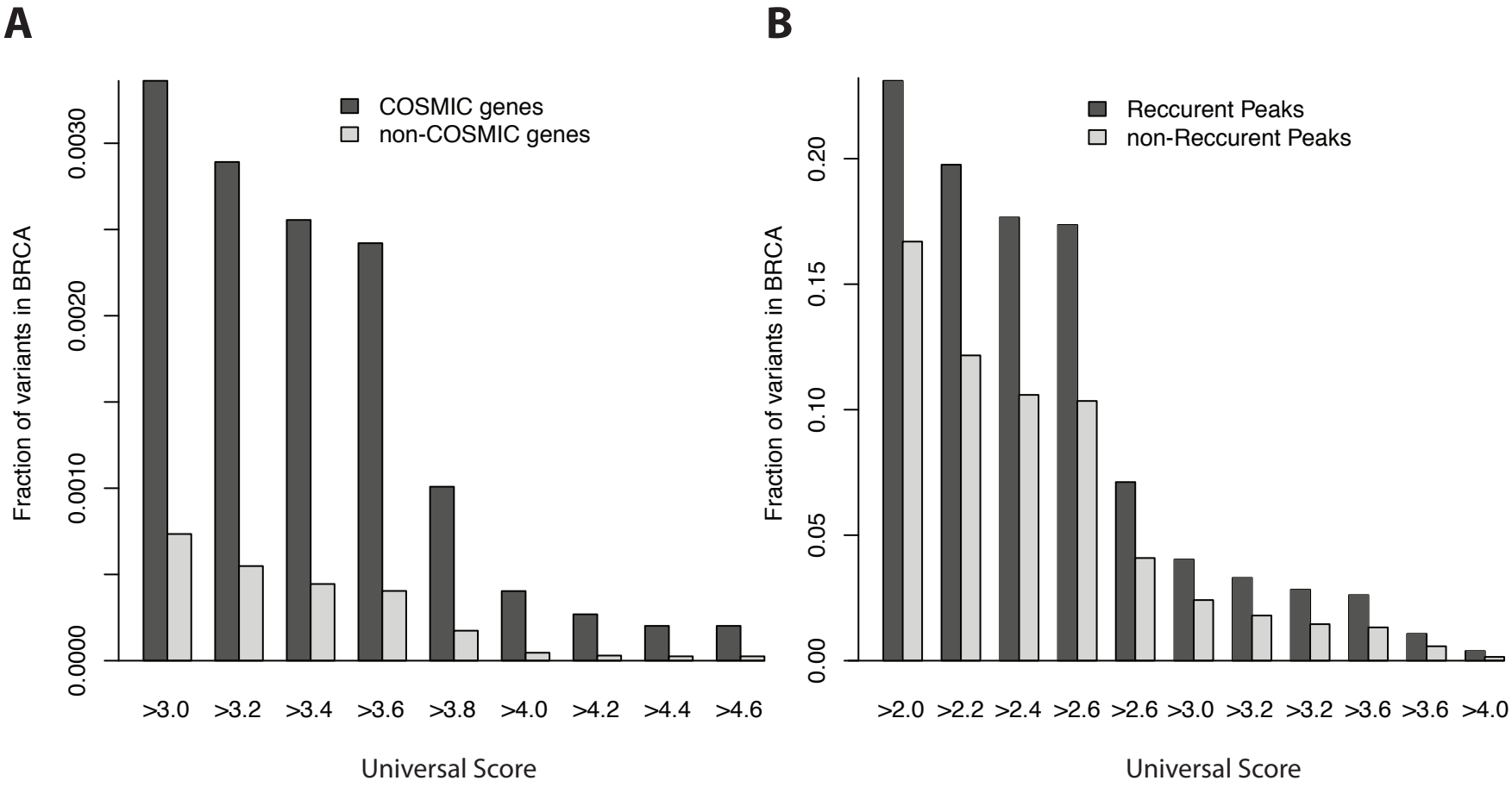
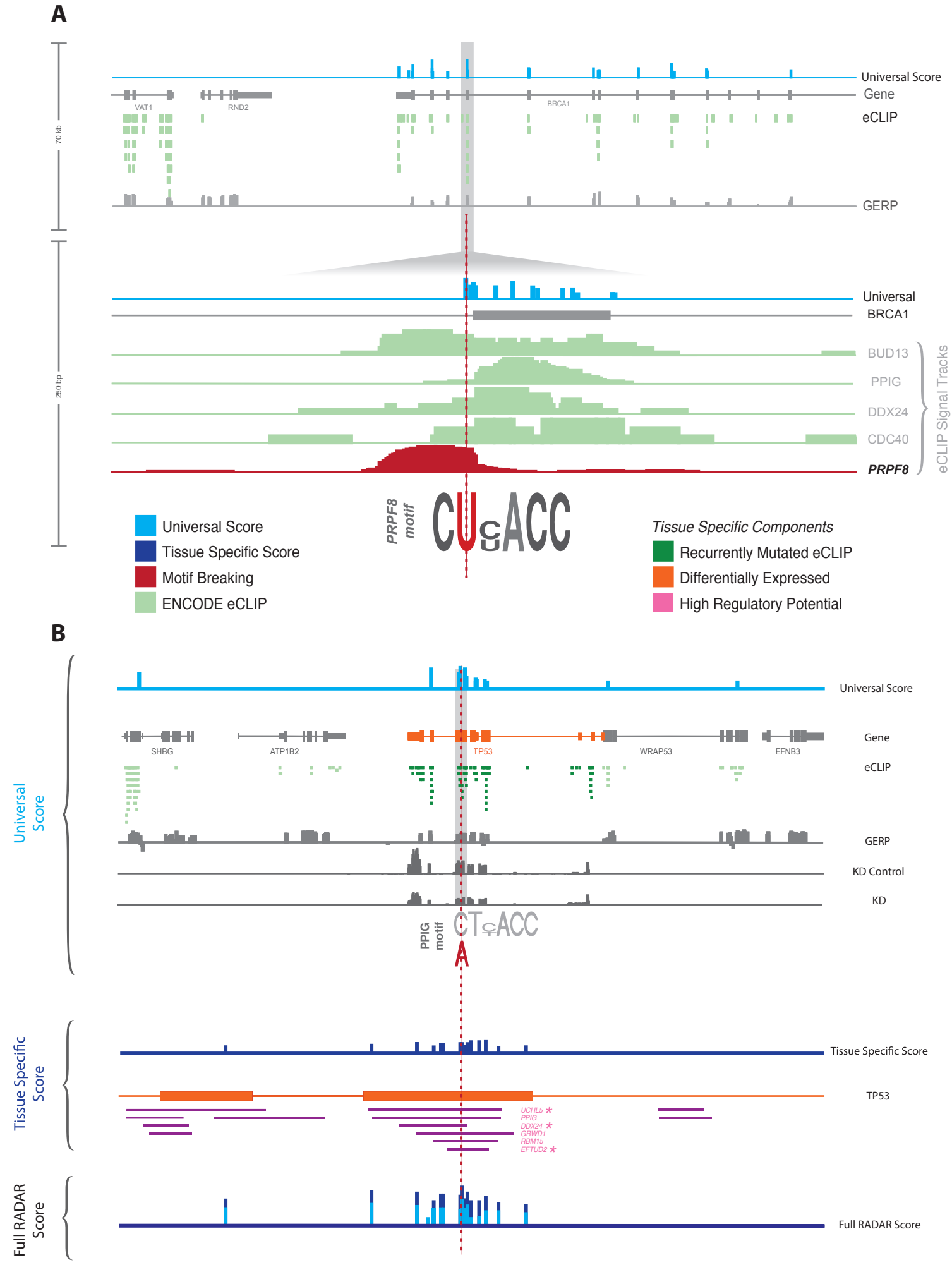

A
COSMIC genes
non-COSMIC genes
Fraction of variants in BRCA
0.0000
0.0010
0.0020
0.0030
>3.0
>3.2
>3.4
>3.6
>3.8
>4.0
>4.2
>4.4
>4.6
Universal Score
B
Reccurent Peaks
non-Reccurent Peaks
Fraction of variants in BRCA
0.00
0.05
0.10
0.15
0.20
>2.0
>2.2
>2.4
>2.6
>2.6
>3.0
>3.2
>3.2
>3.6
>3.6
>4.0
Universal Score

Figure 6


A
70 kb
250 bp
Universal Score
Gene
eCLIP
GERP
VAT1
RND2
BRCA1
Universal
BRCA1
BUD13
PPIG
DDX24
CDC40
PRPF8
eCLIP Signal Tracks
PRPF8 motif
Universal Score
Tissue Specific Score
Motif Breaking
ENCODE eCLIP
Tissue Specific Components
Recurrently Mutated eCLIP
Differentially Expressed
High Regulatory Potential
B
Universal Score
Gene
eCLIP
GERP
KD Control
KD
SHBG
ATP1B2
TP53
WRAP53
EFNB3
PPIG motif
Tissue Specific Score
TP53
UCHL5
PPIG
DDX24
GRWD1
RBM15
EFTUD2
Full RADAR Score