Legends to supplementary figures

Figure S1. Characterization of the Zeb2-V5 mESC lines. (**A**) Schematic overview of the design strategy, showing also primers used for genotyping and detection of the *Zeb2* tagged allele; (**B**) Genotyping results (selected part) showing the "heterozygous" band (red arrow) present in mouse (m)ESC clone 2BE3. The lower band represents the wild-type (WT) *Zeb2* allele; (**C**) Zeb2-V5 specific PCR showing the presence of the tagged allele only in clone 2BE3, and not in WT genomic DNA material. Clone 1BD4 was also used as negative control; (**D**) Sanger DNA-sequencing results of clone 2BE3 showing the alignment with the Zeb2V5 sequence designed *in silico*. The last three amino acids (i.e. GME) of WT Zeb2 have been changed to DK to remove a PAM (protospacer adjacent motif) sequence and mutate the remaining PAM sequence, the target of the gRNA, to avoid multiple cutting by Cas9. An *Eco*RI site was inserted between the Flag- and the V5-coding sequences to facilitate the screening, and an artificial STOP codon (indicated by *) was added downstream of the V5-coding sequence.

Figure S2. Cross-reference of Zeb2-V5-bound protein-coding genes and transcriptome of differentiating mouse (m)ESCs. (**A**) Volcano plots showing the distribution of Zeb2-bound genes in the whole transcriptomes of day (D) D4, D6 and D8 neural differentiated mESCs. Red dots depict the genes bound by Zeb2 and grey crosses those not bound; (**B**) About 11 to 14% of the up or down-regulated genes are bound by Zeb2 (red part of the bars) at D4-D8; (**C**) Of the 1,952 genes bound by Zeb2 and found being differentially expressed genes (DEGs) during mESC differentiation, 335 are in common among the three considered time points. *Zeb2* itself is among these common genes and its mRNA expression increases during differentiation (yellow dot). The volcano plots show the distribution of the common genes during differentiation; (**D**) Time point specific DEGs depicted as volcano plots. The grey arrow in D8 panel indicates *Tcf4*.

Figure S3. Top-10 transcription factor (TF) motifs enrichment at peak present in -10/+10 kb from the TSS of up or down-regulated genes. Scatter plots represent the top-10 TF motifs found in the peaks present in the -10/+10 kb from the TSS range in genes which are up or down-regulated during mouse ESC neural differentiation (panel **A** and **B**, respectively). Each dot represents a peak. Enrichment analysis was performed using UniBind (see Materials and Methods section).

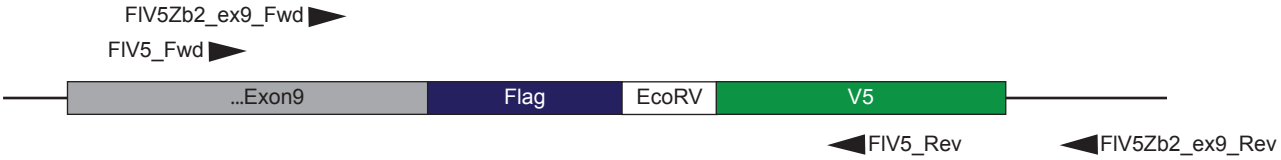
Figure S4. Expression of genes resulting from the meta-analysis of Zeb2-bound genes versus three independent RNA-seq datasets. The heatmap visualizes the log2FoldChange of the resulting genes during mouse (m)ESC differentiation. RNA-seq datasets where the genes have been found to be differentially expressed are annotated as green (Nkx2.1-Cre|Zeb2 mouse model), purple (Zeb2 knockdown (KD) in mESCs) or yellow (Gsh2-Cre|Zeb2 mouse model) squares.

Figure S5. Genotyping of *Zeb2*$^{\Delta P/\Delta P}$ mouse (m)ESCs. (**A**) Schematic overview of the binding sites for the primers (listed in Table 1) used for genotyping; (**B**) PCR confirming the proper deletion of the selected region in ESC clone E9; (**C**) Further PCR showing the difference in fragment length for regions amplified with different sets of primer and discriminating between WT, heterozygous (*Zeb2*$^{\Delta P/+}$) and homozygous (*Zeb2*$^{\Delta P/\Delta P}$) deletion ESC clones; (**D**) Sanger sequencing showing the correct deletion of the selected region in the *Zeb2*$^{\Delta P/\Delta P}$ clone.

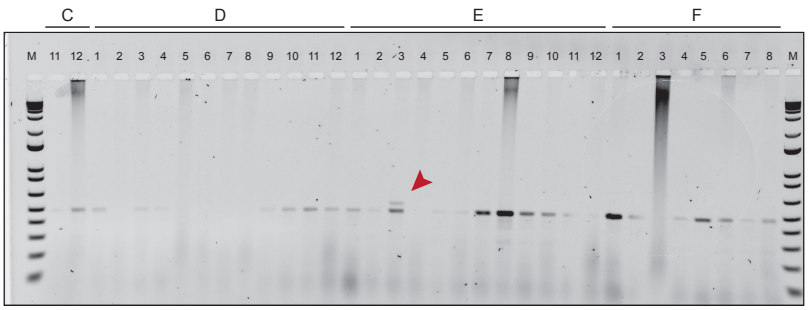
Figure S6. Distribution of Zeb binding sites (performed as single E-boxes; for details, see main text) **and TGFβ/BMP-activated (phospho-)Smad and Smad4 binding elements within the identified Zeb2-bound regions for primary target genes.** (**A**) Grey triangles represent distance from the transcription start site (TSS), while blue triangles depict the TSS of the individual genes. Direction of the arrow shows also whether the gene is transcribed on the + or – strand; (**B**) Motif analysis identifies the binding motif for Zeb2 as CACCTG. Analysis of the distribution and motif analysis were performed using the Jasper database with a confidence score >80% (https://jaspar.genereg.net).

Figure S7. Gene-to-Disease association of the D8 DEGs bound by Zeb2. The gene names of the differentially expressed genes (DEGs) expressed at D8 of ESC neural differentiation were initially subjected to a mouse-to-human gene conversion using biomaRt package for R. After that, gene-to-disease association analysis was performed using disgenet2r package and API dependency, with an association score >30% to retrieve Gene to Disease association from DisGeNET (https://www.disgenet.org/disgenet2r).

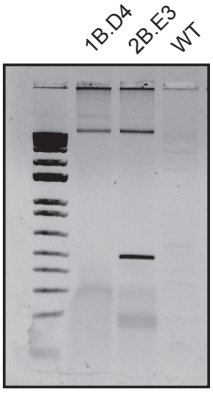
A.


FIV5Zb2_ex9_Fwd
FIV5_Fwd
...Exon9
Flag
EcoRV
V5
FIV5_Rev
FIV5Zb2_ex9_Rev

B.

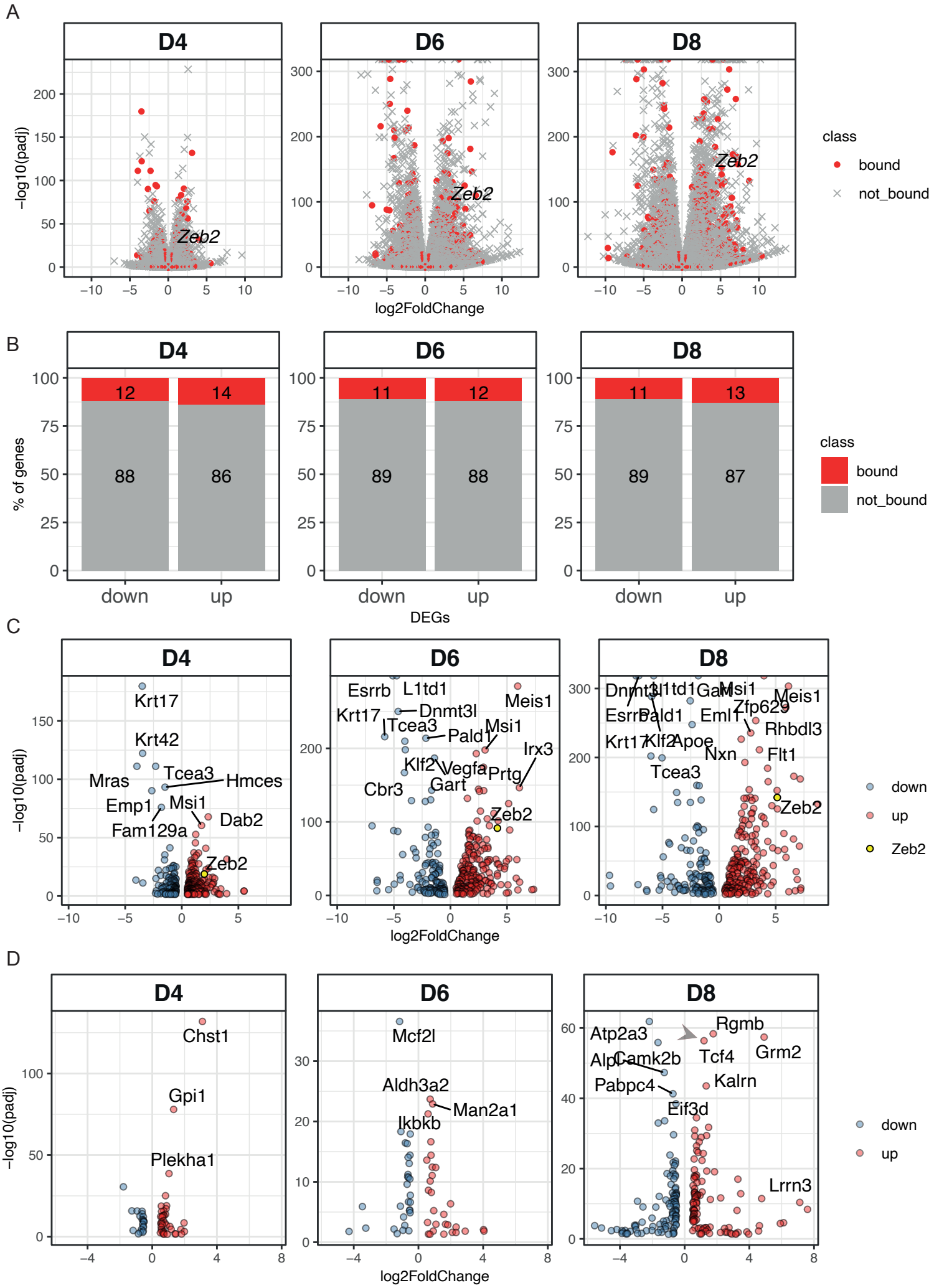

C
D
E
F
M 11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 M

C.

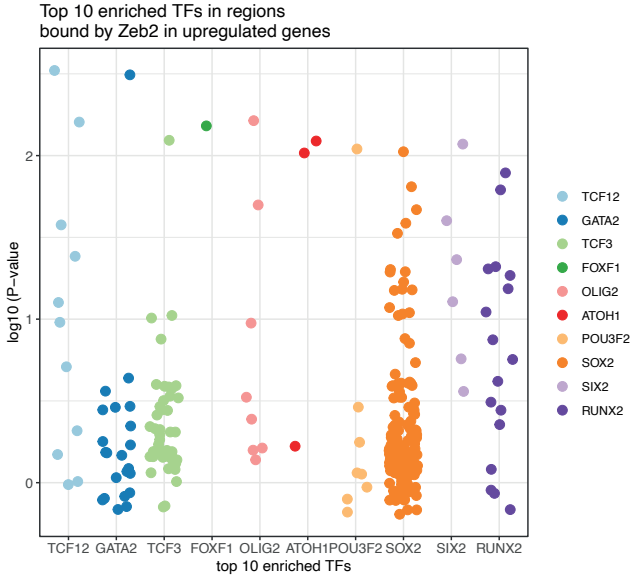

1B.D4
2B.E3
WT

D.

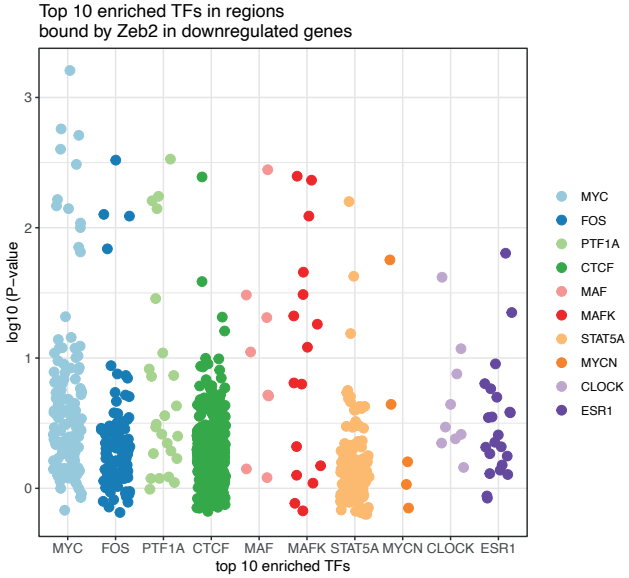

HSMDDSSEDGKMETKSDHEEDNMEDDKDYKDDDDKDIGKPIPNPLLGLDST*
Zeb2FlagV5 sequence
2BE3 alignment
EcoRV
BsaI
SnaBI
...Exon9
Flag
EcoRV
V5

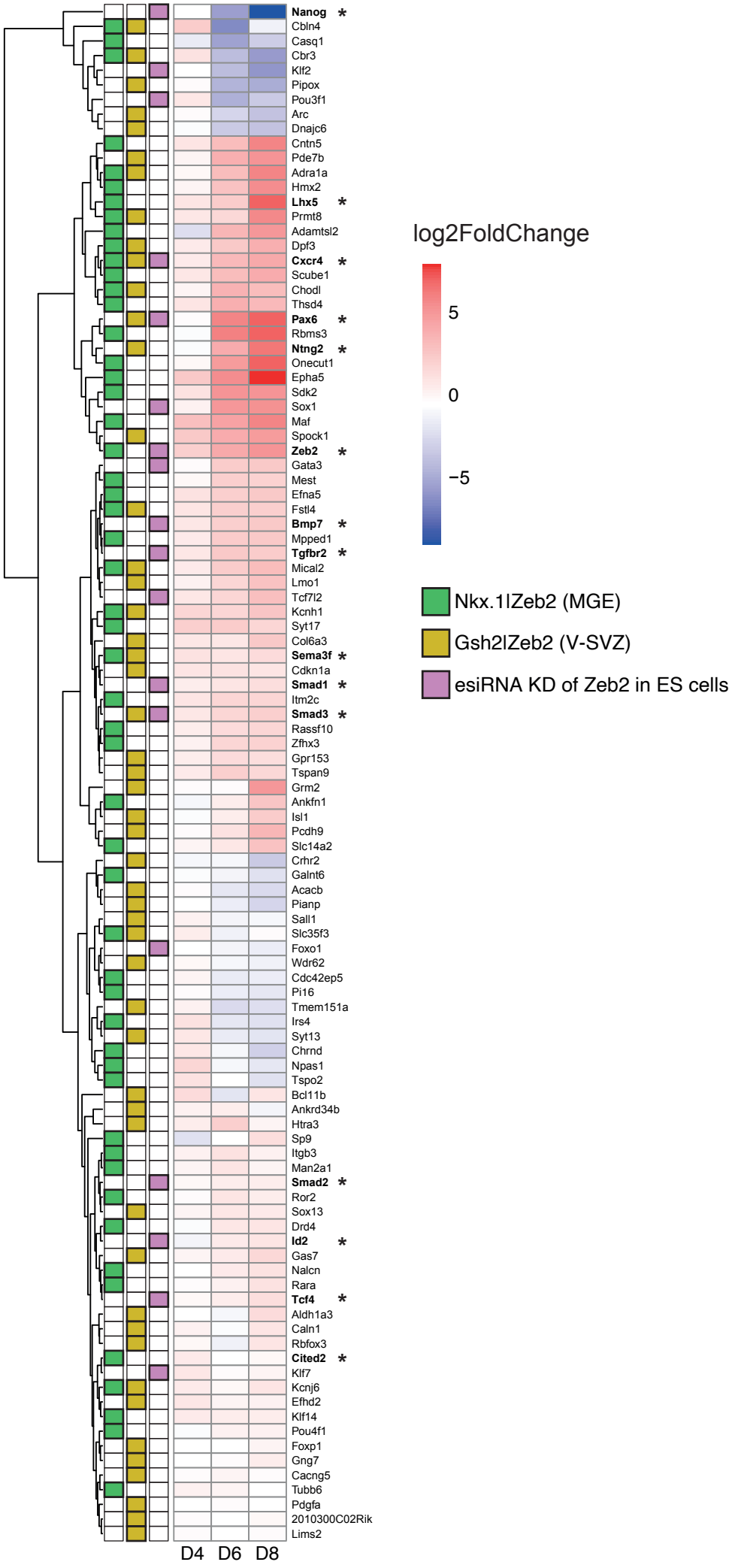

A
D4
D6
D8
−log10(padj)
log2FoldChange
Zeb2
class
bound
not_bound
B
% of genes
DEGs
down
up
12
14
88
86
11
12
89
88
11
13
89
87
C
Krt17
Krt42
Mras
Tcea3
Hmces
Emp1
Msi1
Dab2
Fam129a
Esrrb
L1td1
Dnmt3l
Meis1
Tcea3
Pald1
Irx3
Klf2
Vegfa
Prtg
Cbr3
Gart
Eml1
Zfp629
Rhbdl3
Klf2
Apoe
Nxn
Flt1
down
up
Zeb2
D
Chst1
Gpi1
Plekha1
Mcf2l
Aldh3a2
Man2a1
Ikbkb
Atp2a3
Rgmb
Alp
Camk2b
Tcf4
Grm2
Pabpc4
Kalrn
Eif3d
Lrrn3

A

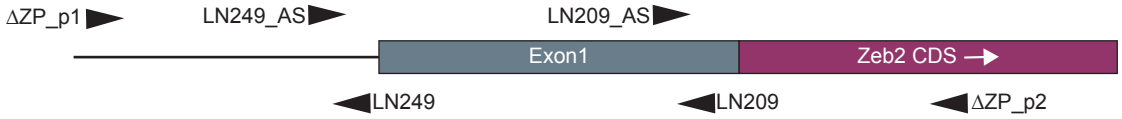

Top 10 enriched TFs in regions
bound by Zeb2 in upregulated genes
log10 (P−value
0
1
2
TCF12 GATA2 TCF3 FOXF1 OLIG2 ATOH1 POU3F2 SOX2 SIX2 RUNX2
top 10 enriched TFs
TCF12
GATA2
TCF3
FOXF1
OLIG2
ATOH1
POU3F2
SOX2
SIX2
RUNX2

B

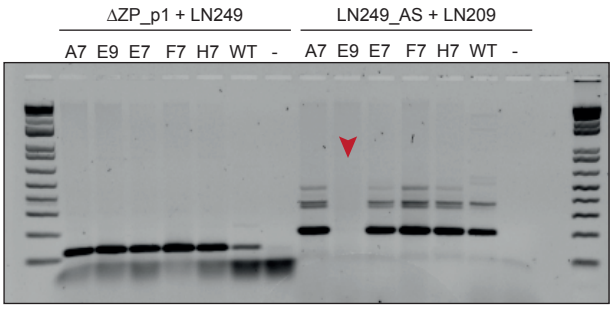

Top 10 enriched TFs in regions
bound by Zeb2 in downregulated genes
log10 (P−value
0
1
2
3
MYC FOS PTF1A CTCF MAF MAFK STAT5A MYCN CLOCK ESR1
top 10 enriched TFs
MYC
FOS
PTF1A
CTCF
MAF
MAFK
STAT5A
MYCN
CLOCK
ESR1


Nanog *
Cbln4
Casq1
Cbr3
Klf2
Pipox
Pou3f1
Arc
Dnajc6
Cntn5
Pde7b
Adra1a
Hmx2
Lhx5 *
Prmt8
Adamtsl2
Dpf3
Cxcr4 *
Scube1
Chodl
Thsd4
Pax6 *
Rbms3
Ntng2 *
Onecut1
Epha5
Sdk2
Sox1
Maf
Spock1
Zeb2 *
Gata3
Mest
Efna5
Fstl4
Bmp7 *
Mpped1
Tgfbr2 *
Mical2
Lmo1
Tcf7l2
Kcnh1
Syt17
Col6a3
Sema3f *
Cdkn1a
Smad1 *
Itm2c
Smad3 *
Rassf10
Zfhx3
Gpr153
Tspan9
Grm2
Ankfn1
Isl1
Pcdh9
Slc14a2
Crhr2
Galnt6
Acacb
Pianp
Sall1
Slc35f3
Foxo1
Wdr62
Cdc42ep5
Pi16
Tmem151a
Irs4
Syt13
Chrnd
Npas1
Tspo2
Bcl11b
Ankrd34b
Htra3
Sp9
Itgb3
Man2a1
Smad2 *
Ror2
Sox13
Drd4
Id2 *
Gas7
Nalcn
Rara
Tcf4 *
Aldh1a3
Caln1
Rbfox3
Cited2 *
Klf7
Kcnj6
Efhd2
Klf14
Pou4f1
Foxp1
Gng7
Cacng5
Tubb6
Pdgfa
2010300C02Rik
Lims2
D4 D6 D8
log2FoldChange
5
0
−5
Nkx.1lZeb2 (MGE)
Gsh2lZeb2 (V-SVZ)
esiRNA KD of Zeb2 in ES cells

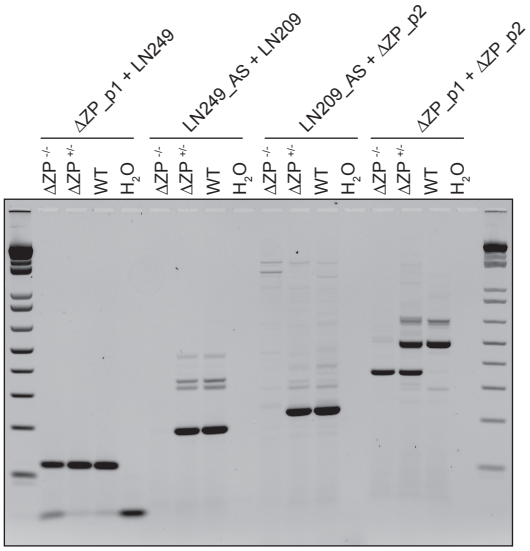
A.


ΔZP_p1
LN249_AS
LN209_AS
Exon1
Zeb2 CDS
LN249
LN209
ΔZP_p2

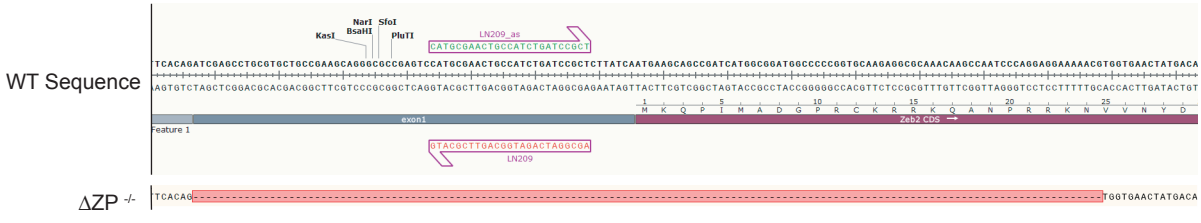
B.


ΔZP_p1 + LN249
LN249_AS + LN209
A7 E9 E7 F7 H7 WT -
A7 E9 E7 F7 H7 WT -

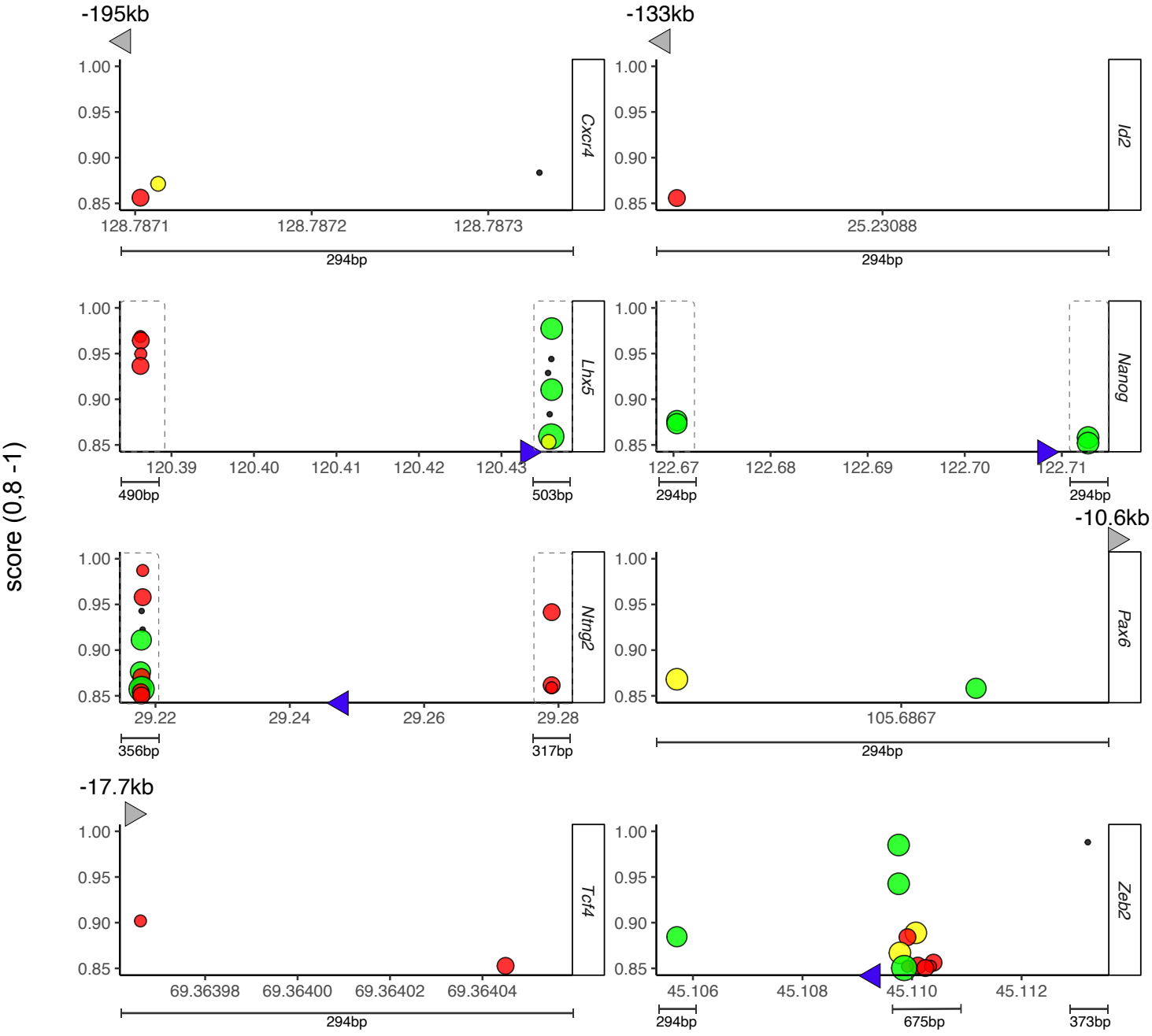
C.


ΔZP_p1 + LN249
LN249_AS + LN209
LN209_AS + ΔZP_p2
ΔZP_p1 + ΔZP_p2
ΔZP -/-
ΔZP +/-
WT
H2O

D.


WT Sequence
ΔZP -/-
NarI
SfoI
KasI
BsaHI
PluTI
LN209_as
CATGCGAACTGCCATCTGATCCGCT
exon1
Zeb2 CDS
Feature 1
GTACGCTTGACGGTAGACTAGGCGA
LN209

A



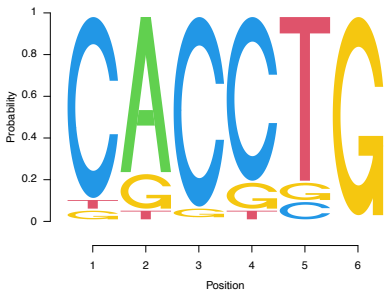
Zeb1/Zeb2

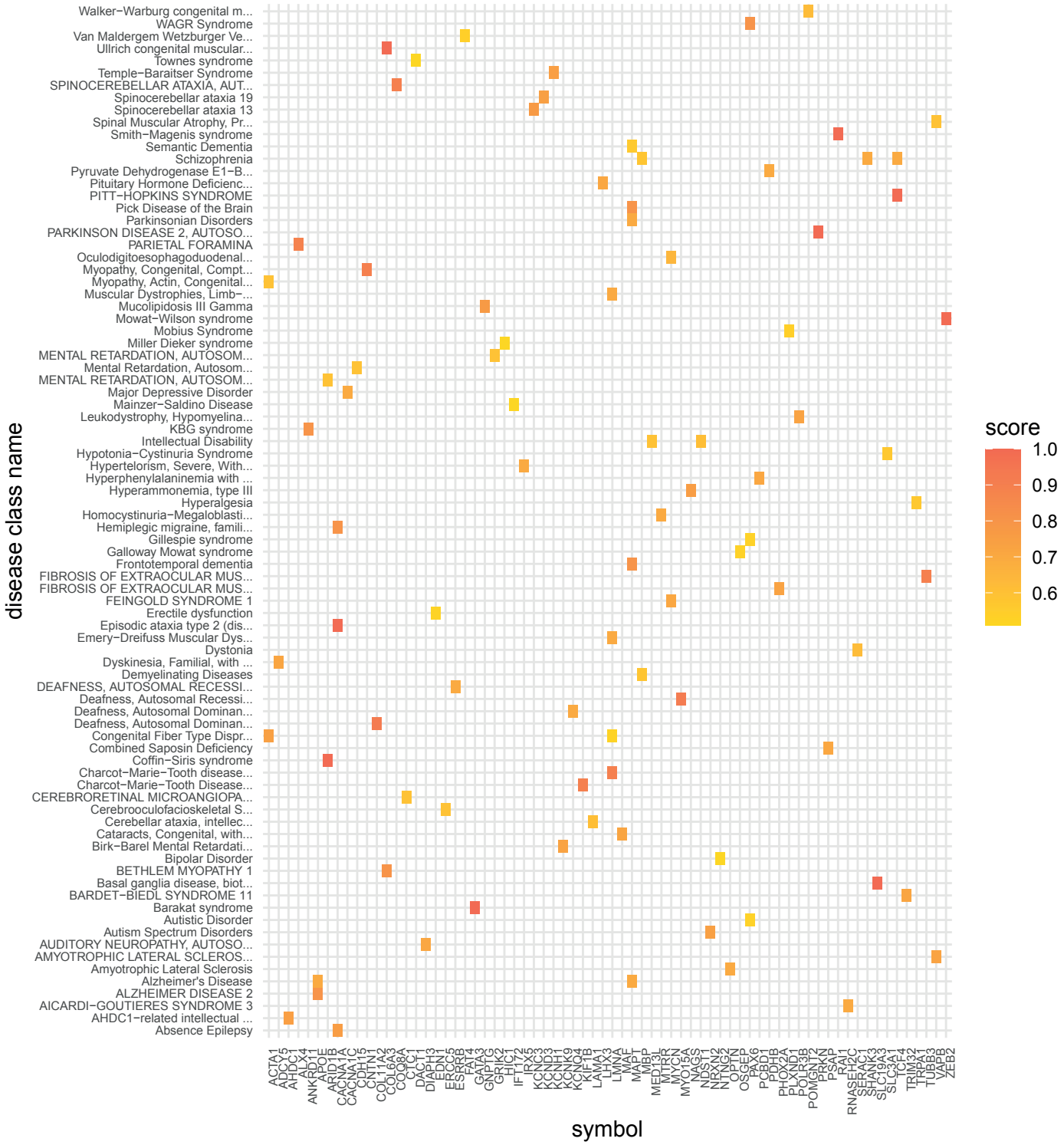
TGFβ R-Smads

BMP R-Smads

Smad4

B




disease class name
Walker−Warburg congenital m...
WAGR Syndrome
Van Maldergem Wetzburger Ve...
Ullrich congenital muscular...
Townes syndrome
Temple−Baraitser Syndrome
SPINOCEREBELLAR ATAXIA, AUT...
Spinocerebellar ataxia 19
Spinocerebellar ataxia 13
Spinal Muscular Atrophy, Pr...
Smith−Magenis syndrome
Semantic Dementia
Schizophrenia
Pyruvate Dehydrogenase E1−B...
Pituitary Hormone Deficienc...
PITT−HOPKINS SYNDROME
Pick Disease of the Brain
Parkinsonian Disorders
PARKINSON DISEASE 2, AUTOSO...
PARIETAL FORAMINA
Oculodigitoesophagoduodenal...
Myopathy, Congenital, Compt...
Myopathy, Actin, Congenital...
Muscular Dystrophies, Limb−...
Mucolipidosis III Gamma
Mowat−Wilson syndrome
Mobius Syndrome
Miller Dieker syndrome
MENTAL RETARDATION, AUTOSOM...
Mental Retardation, Autosom...
MENTAL RETARDATION, AUTOSOM...
Major Depressive Disorder
Mainzer−Saldino Disease
Leukodystrophy, Hypomyelina...
KBG syndrome
Intellectual Disability
Hypotonia−Cystinuria Syndrome
Hypertelorism, Severe, With...
Hyperphenylalaninemia with ...
Hyperammonemia, type III
Hyperalgesia
Homocystinuria−Megaloblasti...
Hemiplegic migraine, famili...
Gillespie syndrome
Galloway Mowat syndrome
Frontotemporal dementia
FIBROSIS OF EXTRAOCULAR MUS...
FIBROSIS OF EXTRAOCULAR MUS...
FEINGOLD SYNDROME 1
Erectile dysfunction
Episodic ataxia type 2 (dis...
Emery−Dreifuss Muscular Dys...
Dystonia
Dyskinesia, Familial, with ...
Demyelinating Diseases
DEAFNESS, AUTOSOMAL RECESSI...
Deafness, Autosomal Recessi...
Deafness, Autosomal Dominan...
Deafness, Autosomal Dominan...
Congenital Fiber Type Dispr...
Combined Saposin Deficiency
Coffin−Siris syndrome
Charcot−Marie−Tooth disease...
Charcot−Marie−Tooth Disease...
CEREBRORETINAL MICROANGIOPA...
Cerebrooculofacioskeletal S...
Cerebellar ataxia, intellec...
Cataracts, Congenital, with...
Birk−Barel Mental Retardati...
Bipolar Disorder
BETHLEM MYOPATHY 1
Basal ganglia disease, biot...
BARDET−BIEDL SYNDROME 11
Barakat syndrome
Autistic Disorder
Autism Spectrum Disorders
AUDITORY NEUROPATHY, AUTOSO...
AMYOTROPHIC LATERAL SCLEROS...
Amyotrophic Lateral Sclerosis
Alzheimer's Disease
ALZHEIMER DISEASE 2
AICARDI−GOUTIERES SYNDROME 3
AHDC1−related intellectual ...
Absence Epilepsy
ACTA1
ADCY5
AHDC1
ALX4
ANKRD11
APOE
APP
CACNA1A
CACNB4
CDH15
CNTN1
COL11A2
COL6A3
COQ8A
CTC1
DAG1
DIAPH3
EDN1
ESRRB
EST4
FAT4
GATA3
GNPTG
GRIK2
HJC1
IFT172
IRX5
KCNC3
KCND3
KCNH1
KCNK9
KCNQ4
KIF1B
LAMA1
LHX3
LMNA
MAF
MAPT
MBP
MED13L
MTRR
MYO7A
NAGS
NDST1
NRXN2
NTNG2
OPTN
OSGEP
PAX6
PCBD1
PDHB
PHOX2A
PLXNB1
POLR3B
POMGNT2
PRKN
PSAP
RAI1
RNASEH2C
SERAC1
SHANK3
SLC19A3
SLC2A1
TCF4
TRIM32
TUBA1A
TUBB3
VAPB
ZEB2
symbol
score
1.0
0.9
0.8
0.7
0.6