Supplemental Material

ZNF91 deletion in human embryonic stem cells leads to ectopic activation of SVA retrotransposons and upregulation of KRAB zinc finger gene clusters

Haring et al., 2021

Supplemental_Fig_S1

Supplemental_Fig_S2

Supplemental_Fig_S3

Supplemental_Fig_S4

Supplemental_Fig_S5

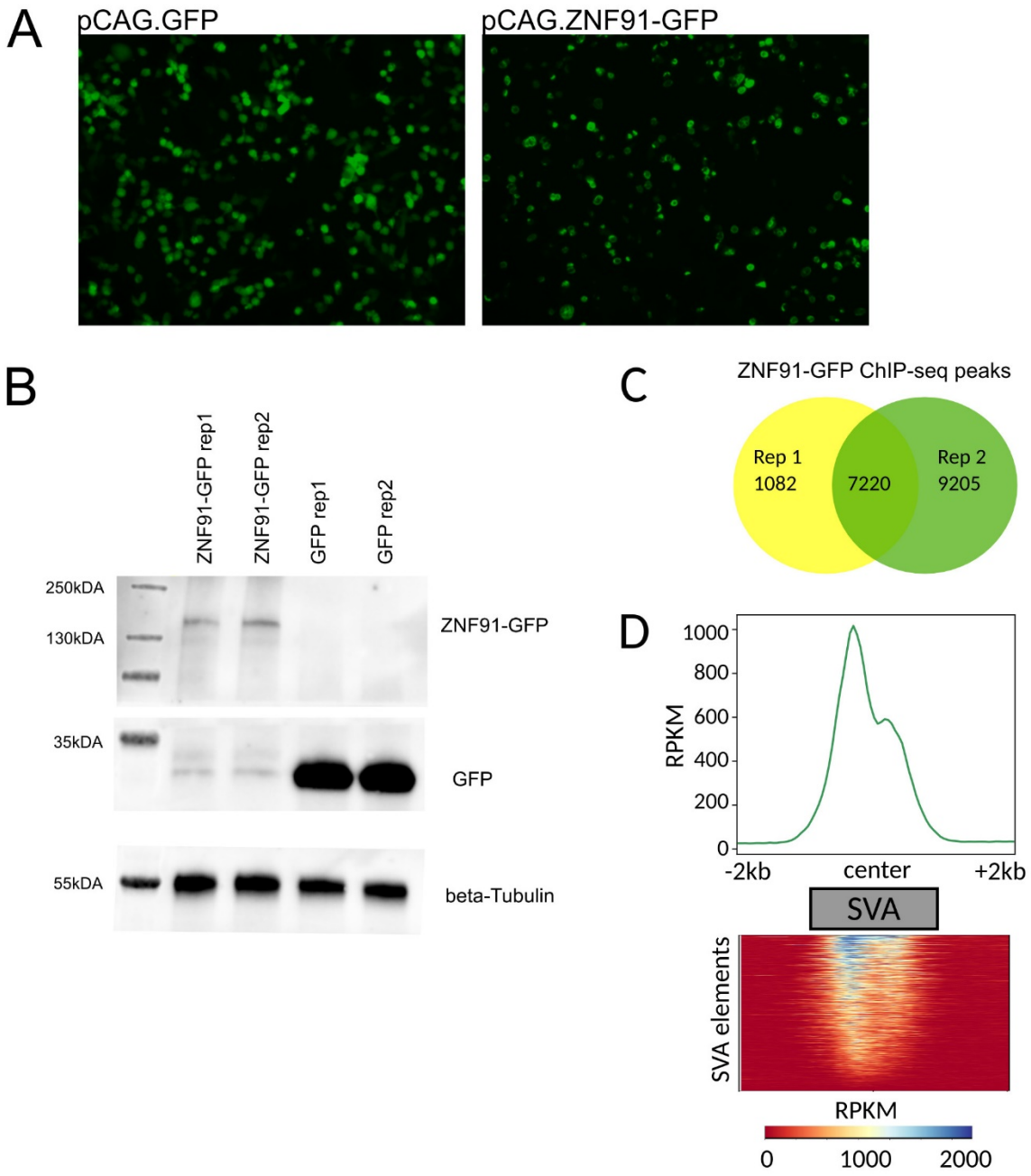
Supplemental_Table_S1

Supplemental_Table_S2

Supplemental_Table_S3

Supplemental_File_1

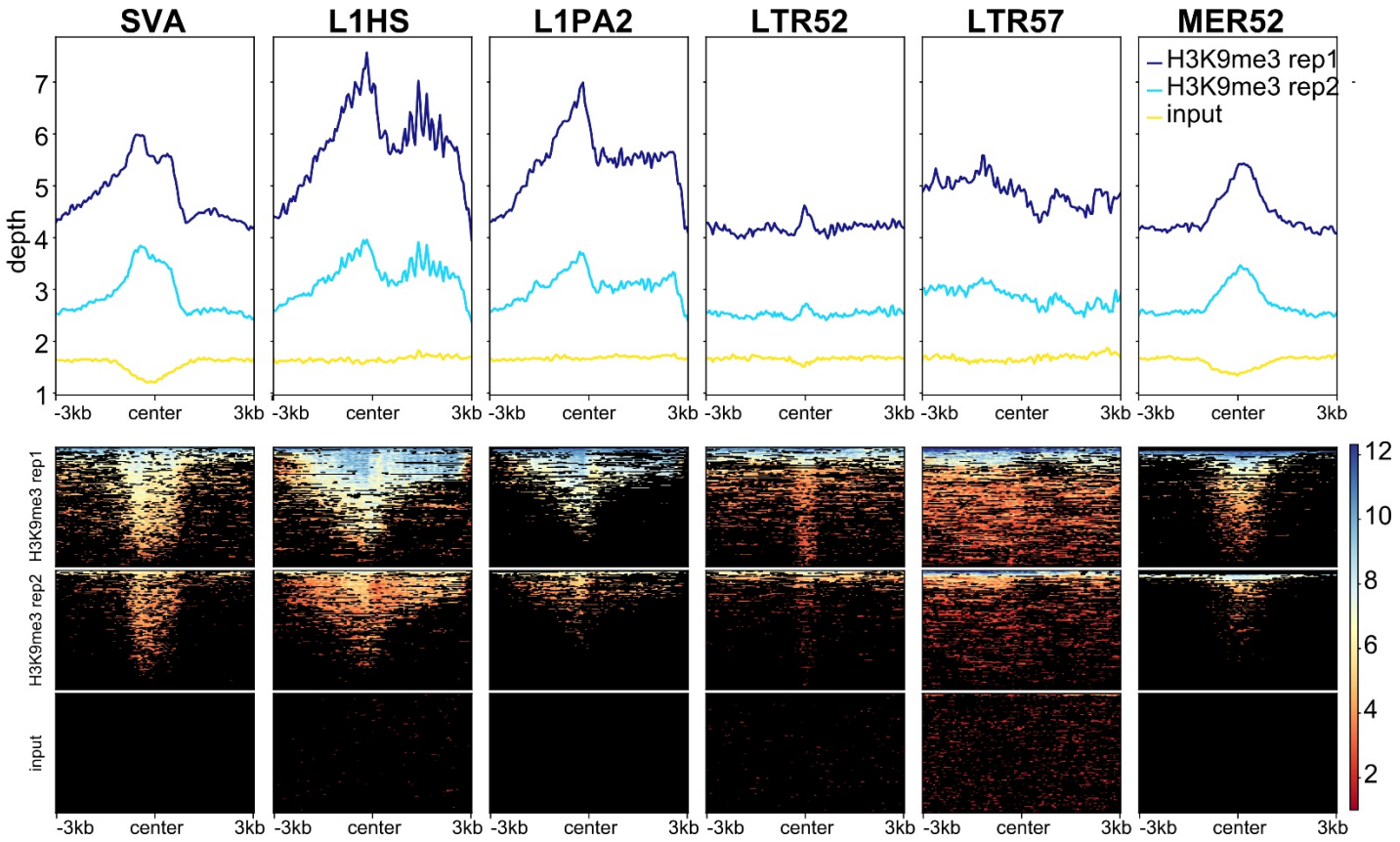
Supplemental_Fig_S1



Supplemental Figure 1. ZNF91-GFP ChIP.

A) HEK293 cells transfected with pCAG.GFP (left) and pCAG.ZNF91-GFP (right), 48 hours post-transfection. B) Westernblot for GFP and beta-Tubulin lysates from cells displayed in (A). ZNF91-GFP is detected at 166 kDA, GFP at 29 kDA, and beta-Tubulin at 55 kDA. C) Venn diagram showing number of ZNF91-GFP peaks derived from replicate 1 and 2 through MACS peak calling. D) Heatmap and profile plot of ZNF91 ChIP showing coverage in RPKM of the mean of rep 1 and rep 2 with SVA center as a reference point. Average signal across all SVA elements is displayed.

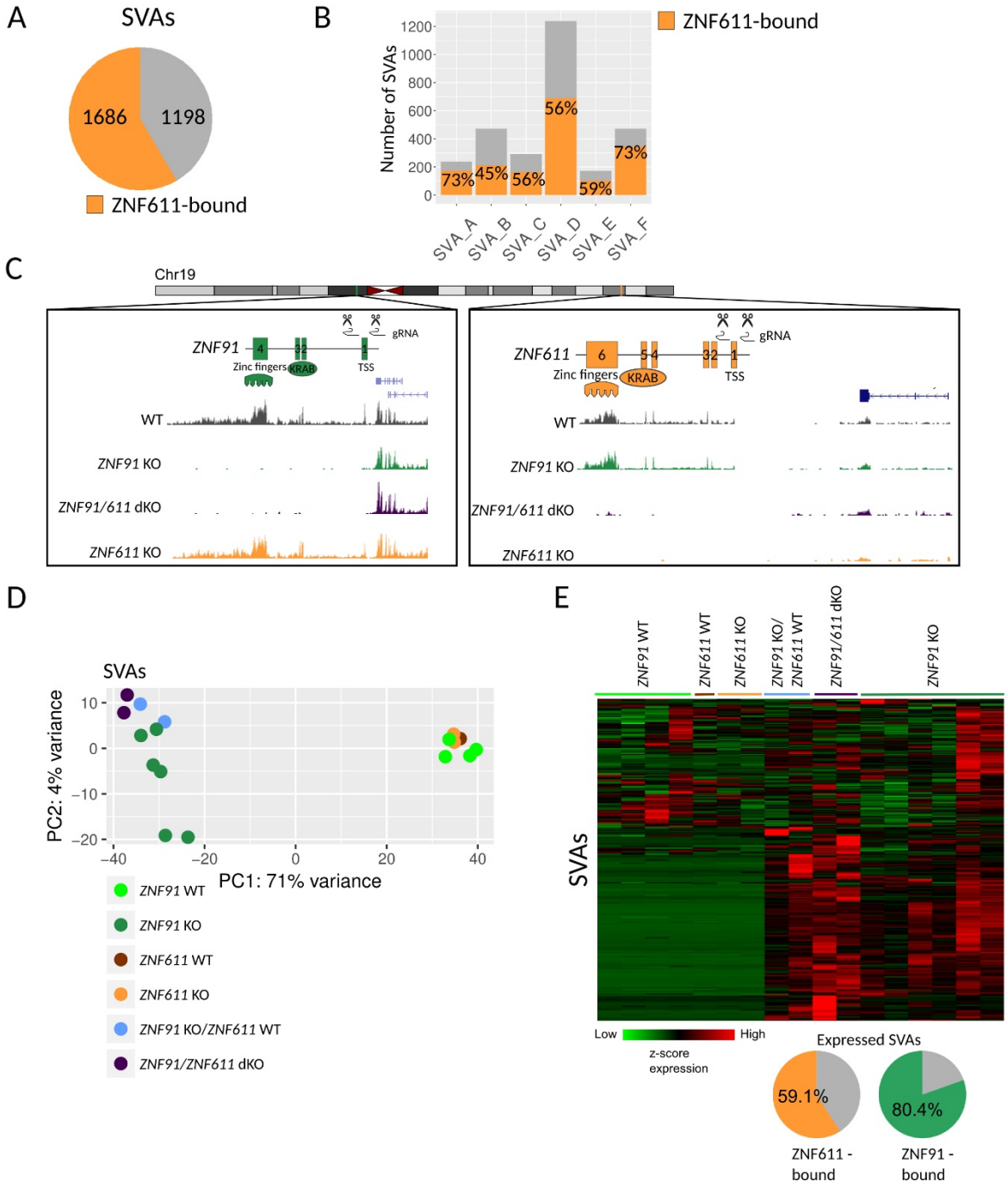
Supplemental_Fig_S2



Supplemental Figure 2. Profile plots of H3K9me3 on several TE classes in hESCs.

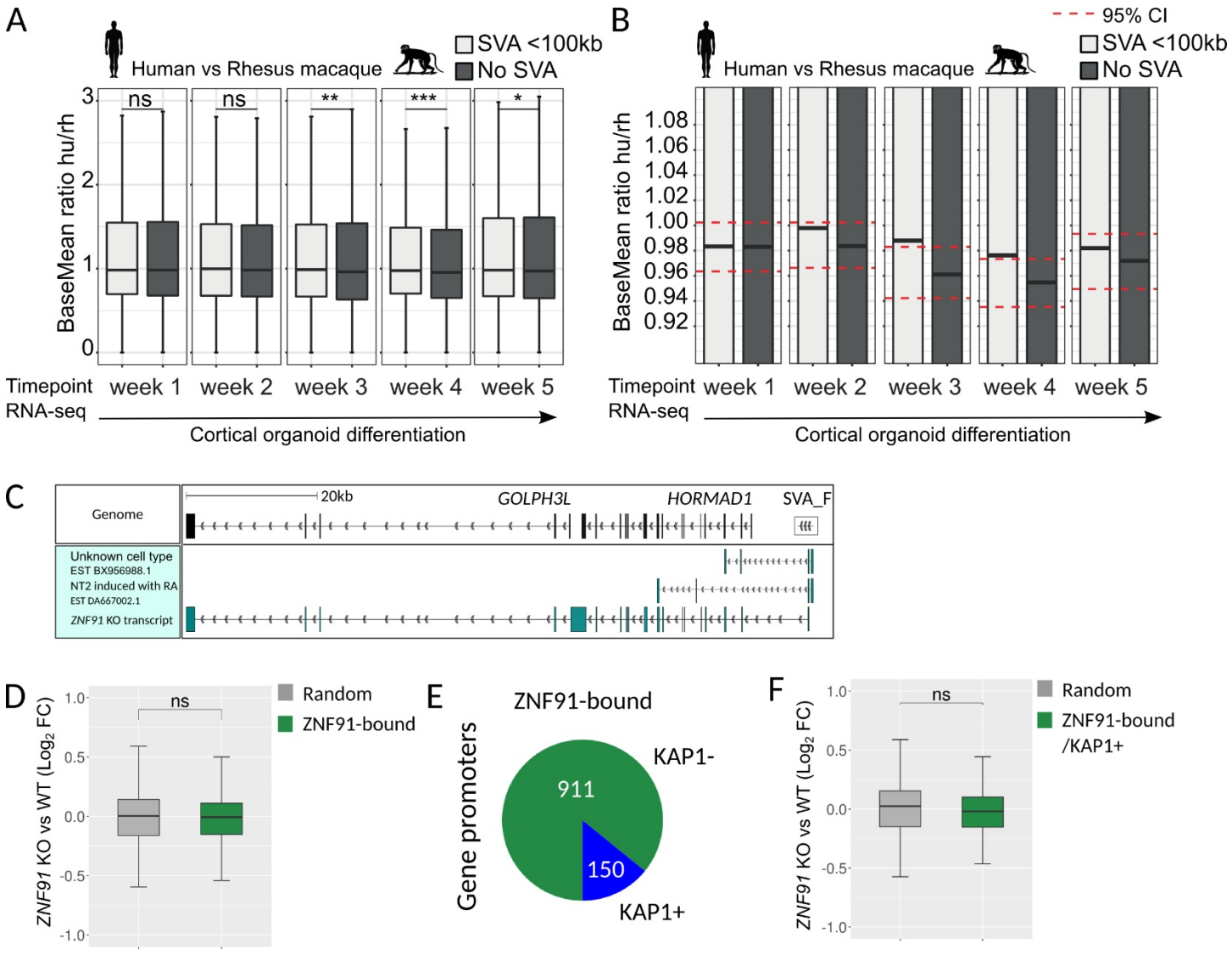
Heatmaps and profile plots of H3K9me3 signal showing coverage across several TE subclasses, using the center of the element as reference and extending 3kb in both directions. Two replicates of H3K9me3 ChIP and one input sample of WT hESCs were included.

Supplemental_Fig_S3



Supplemental Figure 3. ZNF611 appears non-essential for SVA repression in hESCs. A) Pie chart showing the number of SVA elements that are bound (orange, 1686) and not bound (grey, 1198) by ZNF611. B) Frequency plot showing fraction of SVA subclasses bound by ZNF611 (orange). C) Schematic overview of *ZNF91* and *ZNF611* genes. gRNAs targeting the area around the TSSs were used to generate *ZNF91* ko, *ZNF611* KO and *ZNF91*/*ZNF611* dko hESCs using CRISPR-Cas9. Coverage tracks of RNA-seq display (lack of) expression of *ZNF91* and *ZNF611* in hESCs. RNA-seq coverage tracks were scaled using DESeq2 scaling factors. ChIP-seq coverage tracks were scaled based on control regions. D) Principal component analysis including all SVA transcripts in hESC lines with indicated genotype. E) Heatmap showing hierarchical clustering of SVA elements that are expressed in any sample (expressed defined as >20 reads) using Z-score of normalized counts of RNA-seq data of *ZNF91* and *ZNF611* WT and KO hESC lines and *ZNF91*/*ZNF611* dko lines. Pie charts showing fractions of expressed SVA elements that are bound by ZNF611 or ZNF91 in HEK293(T).

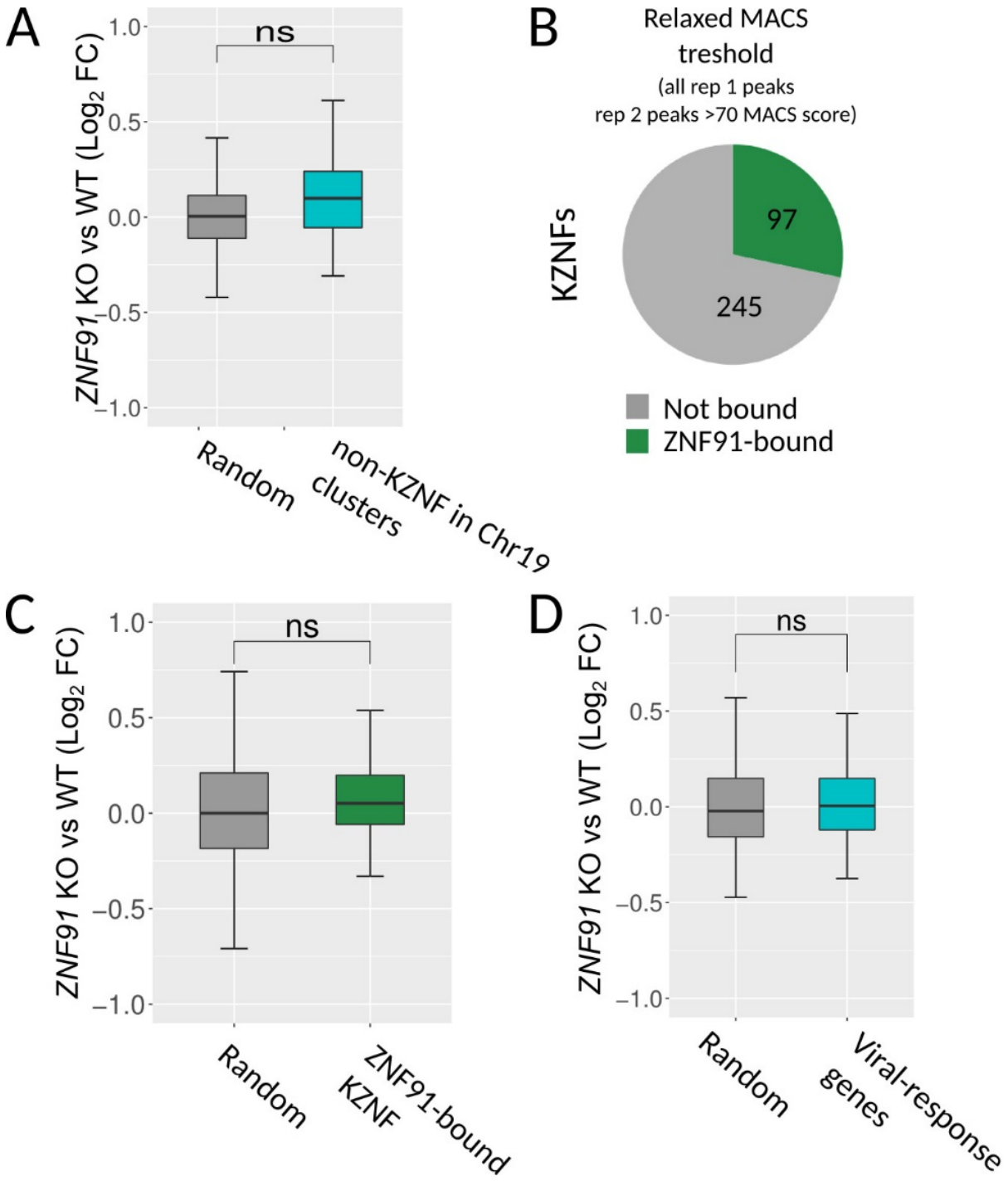
Supplemental_Fig_S4



Supplemental Figure 4. ZNF91/SVA-mediated gene regulation

A-B) Ratio of basemean (DESeq2) expression levels between human and rhesus macaque shown for week 1/5 cortical organoids for transcripts with an SVA within 100kb (light grey, 4507) and without (dark grey, 18596). Boxplots show median (black line), interquartile range (box), and 1.5x interquartile range (whiskers). B) Zoom in on median with 95% CI of the median of 10.000 similar sized groups of random genes without SVA nearby (red). C) *De novo* transcript assembly of RNA-seq data of *ZNF91* KO hESCs indicates that a part of SVA-F becomes spliced into the *HORMAD1* transcript generating the SVA-*HORMAD1-GOLPH3L* fusion transcript. Expressed sequence tags (ESTs) covering the *SVA-HORMAD1* fusion transcript. D) Boxplots showing $\log_2$ fold change of 799 genes bound by ZNF91 at their promoters compared to a size matched random set of genes in *ZNF91ko* versus WT hESCs. E) Pie chart showing the number of gene promoters having KAP1 binding in hESCs (data from Jacobs et al 2014) at the same location as ZNF91 binding was detected in HEK293 cells. F) Boxplots showing the expression $\log_2$ fold change of *ZNF91* KO versus WT hESCs for genes bound by KAP1 and ZNF91 at their promoters (see E) compared to a size matched random set of genes. Only expressed (baseMean>10) were included. For statistical testing an unpaired wilcoxon rank sum test was used. Outliers and individual data points not shown. Wilcoxon rank sum test, *** P<0.001, ** P<0.01, * P<0.05, ns= not significant.

Supplemental_Fig_S5



Supplemental Figure 5. Gene expression changes in *ZNF91* KO hESCs.
A) Boxplot showing log_2 fold change of 34 non-KZNF genes residing in the upregulated KZNF clusters on Chromosome 19 compared to 46 randomly selected genes in *ZNF91 KO* hESCs. B) Pie chart showing in green the number of KZNF genes bound by ZNF91 at their promoter (TSS plus/minus 1000bp) using a relaxed peak calling threshold (all peaks of rep 1 and peaks with a MACS score>70 of rep 2). C) Boxplots showing log_2 fold change of 92 ZNF91-bound KZNF genes compared to a group of 118 randomly selected genes. D) Boxplot showing log_2 fold change of 174 viral response genes (GO term: response to virus) compared to 180 randomly selected genes. For statistical testing an unpaired wilcoxon rank sum test was used (ns= not significant).

Supplemental_Table_S1:

Gene	Direction	SVA	Distance from TSS to SVA	SVA H3K4me3+	SVA H3K27ac+	Putative SVA mediated regulation
AMDHD1	Up	SVA_F	3.3 kb downstream (Intron 1)	Yes	No	Alternative TSS
CAMP	Up	SVA_F	10 kb upstream	Yes	No	Alternative TSS
CLDN4	Up	SVA_D	2.5 kb downstream	Yes	No	Cis regulation
DLEC1	Up	SVA_F	1.4kb downstream (Intron 6)	Yes	Yes	Alternative TSS
GOLGA7B	Up	SVA_D	8kb upstream	yes	Yes	Alternative TSS
HORMAD1	Up	SVA_F	7kb upstream	Yes	Yes	Alternative TSS
LOC100129269	Up	SVA_D	At TSS	Yes	Yes	Alternative TSS
MRGPRX3	Up	SVA_D	At TSS	Yes	Yes	Alternative TSS
PSPC1	Up	SVA_F	2kb downstream (Intron 2)	Yes	Yes	Extended promoter
RNASE10	Up	SVA_E	At TSS	Yes	Yes	Alternative TSS
SYT11	Up	SVA_D	25kb upstream	Yes	No	Cis regulation
THOC5	Up	SVA_C	0.4 kb downstream (Intron 1)	Yes	Yes	Extended promoter
TOR4A	Up	SVA_F	5kb downstream	Yes	No	Cis regulation
ZNF772	Up*	SVA_D	5kb upstream	Yes	No	Cis regulation
TACR3	Up*	SVA_B	3kb upstream	Yes	No	Cis regulation
PLEKHA7	Unchanged	SVA_F	35 kb upstream	Yes	Yes	Alternative TSS
VPS37C	Unchanged	SVA_D	4.5kb upstream	Yes	Yes	Alternative TSS
EML2	Down	SVA_D	2kb upstream	Yes	No	Cis regulation
RNF135	Down	SVA_F	7k.8 kb downstream (Intron 1)	Yes	No	Cis regulation
WBSCR27	Down	SVA_D	5.2kb downstream (Intron 5)	Yes	No	Cis regulation

Supplemental_Table_S1: SVA controlled genes in *ZNF91* KO human embryonic stem cells. Genes that appear to be regulated by an SVA in *ZNF91* KO hESCs. Green indicates significant upregulation (padj <0.05), light green indicates a trend towards upregulation (padj <0.1), grey indicates unchanged expression, and red indicates significant downregulation (padj <0.05). Putative SVA-mediated regulation is based on inspection of RNA coverage, *de novo* transcript assembly, and H3K4me3/H3k27ac ChIP peaks in close vicinity.

Supplemental_Table_S2

Gene	Oligo name	Sequence (5' → 3')
ZNF91	gRNA upstream exon 1	TCGAGACCTGGAAACTCCGGCGG
ZNF91	gRNA downstream exon 1	ACGTCCCGAGAGAGGGAACGGGG
ZNF91	Genotyping primer forward	ATCTAATCAGGGACGCTGGGC
ZNF91	Genotyping primer reverse	ATTCATGAGCCAGCACCTCC
ZNF91	qPCR primer forward	CCAGACCTGATTACTTATCTGG
ZNF91	qPCR primer reverse	ACATTTTTCATATTTTCTCAGTAATAC
ZNF611	gRNA upstream exon 1	AAACAGAGACGCTGGGAGCG
ZNF611	gRNA downstream exon 1	ACATCCCCGTGAGAGCCCGA
ZNF611	Genotyping primer forward	GCGAGGAGTCAGAACAAGGT
ZNF611	Genotyping primer reverse	GCAGGGCTTTGCATTTCACAT
ZNF611	qPCR primer forward	CATAATTGCACCCTCCGACAT
ZNF611	qPCR primer reverse	CATCCAGAGGACAGCCCCTCA

Supplemental_Table_S2: Oligos used for the generation of *ZNF91* and *ZNF611* KO hESCs.

Supplemental_Table _S3

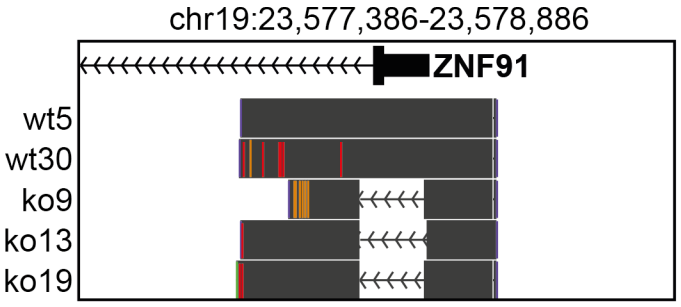
Target sequence	Sequence (5' → 3')
SVAs #1_forward	GTGTACCCAACAGCTCATTG
SVAs #1 reverse	CACGGCAACCATCCGATTTC
SVAs #2 forward	GCCTTGGGATCCTGTTGATC
SVAs #2 reverse	CTTAACGAGCATGCTGCCTTC
LTR12C forward	GTCTCGCTGGCTCAGGAGTG
LTR12C reverse	TGAGCTGTAACACTCACCGC

Supplemental_Table _S3: Oligos used for H3K9me3 ChIP-qPCR.

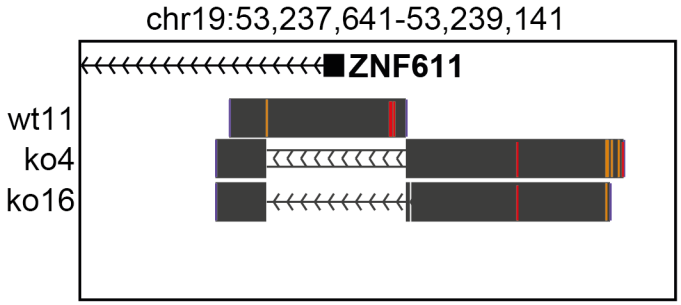
Supplementary_File_1

Sanger Sequencing results of ZNF91 and ZNF611 KO and WT clones

ZNF91 CRISPR KO: WT (wt5; wt30) and KO (ko9; ko13; ko19) clones
BLAST results on ZNF91 locus



ZNF611 CRISPR KO: WT (wt11) and KO (ko4; ko16) clones
BLAST results on ZNF91 locus



Sanger sequencing results

ZNF91 wt#5
>170914-066_I06_1AA4PAA028_E3.ab1 790
GGGTGACGCTACAGCGTCTCTTTCTTATTCAGCCCAGTGTCTGATCACAT
CTTCTGTCACTCAGGGCCTGAGAAGGCGGGGCCTTAAACGTTATCTAATC
AGGGACGCTGGGCTGGGAACTGTCCAATCAGGCACAAAGCTGGAGCGGAC
AGGGCGGCTTCCGGGATGTGGCTGGGCCTTTGTCTCTCGCTGCCGCCGGA
GTTTCCAGGTCTCGACTTCACTGCTCTGTGTCCTCTGCTCCAGGAGGCCC
AGCCTGTGTGGCCCTGTGACCTGCAGGTATTGGAGAGCCACAGCTAAGAT
GCCAGGAACCCCTGGAAGCCTAGAAATGGTGAGACTGCCAGGTGCGACGT
CCCGAGAGAGGGAACGGGGCTCGTTGAAACCGGTGGGATGCGGCTGTGGC
GAGCCTCAGGCCTTCCTTCAGTCAGCTCCACAATCCGTGCCCGGGTTCTC
CTTGCCCAGCTCGGCCTCAGTCCCCTTCAGCGATAAGATGGCGGCTGCGT
TGACAGGCGGGCTCCGGGCGTCCTGTCTCTTCCCTGCCCAGTGACTGTGT
CCTGGCCTGGAGCCCTCTCTGGGCAGCTCTGCTCCCGCAGCGCCGCGTCT
CTCCCAGATTGTGCAGGGATCACGGGAGGGTCGTCAGGGGAGAATCCTGA
CTGGAGGTGCTGCCTCCTTGAATAAAAAAGCGGGGGGGGGGGGGGGGGGG
CCGGGGGGGGGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
GGGGCGGCCCGGGCGGGGGGGGGCCGCGCGGCGCGGGCGG

ZNF91 wt#30
>170914-066_I12_1AA4PAA028_E6.ab1 949
TAGGTGCAAGTCAGCGTCTCTTTCTTATTCAGCCCAGTGTCTGATCACAT
CTTCTGTCACTCAGGGCCTGAGAAGGCGGGGCCTTAAACGTTATCTAATC
AGGGACGCTGGGCTGGGAACTGTCCAATCAGGCACAAAGCTGGAGCGGAC
AGGGCGGCTTCCGGGATGTGGCTGGGCCTTTGTCTCTCGCTGCCGCCGGA
GTTTCCAGGTCTCGACTTCACTGCTCTGTGTCCTCTGCTCCAGGAGGCCC
AGCCTGTGTGGCCCTGTGACCTGCAGGTATTGGAGAGCCACAGCTAAGAT
GCCAGGAACCCCTGGAAGCCTAGAAATGGTGAGACTGCCAGGTGCGACGT
CCCGAGAGAGGGAACGGGGCTCGTTGAAACCGGTGGGATGCGGCTGTGGC
GAGCCTCATGCCTTCCTTCAGTCAGCTCCACAATCCGTGCCCGGGTTCTC
CTTGCCCAGCTCGGCCTCAGTCCCCTTCAGCGATAAGATGGCGGCTGCGT
TGACAGGCGGGCTCCGGGCGTCCTGTCTCTTCCCTGCCCAGTGACTGTGT
CCTGGGCTGGAGACATCTCTGGGCAGCTCTGCTCCCGCAGCGCCGCGTCT
CTCCCAAATTGTGCAGGGATCACGGGAGGGTCGTCAGTGGGAGAATCCTG
ACTGGCGGTGCTGGCCTCTTCATTTCACCTTCAAGAGGGCTCTATTACGA
TCTTTAATATACCCTATGATTTGGCTCCATCGCCAAGAATTGATATCATT
CGCCTCCCTATTCCGAGAAGAGAAAACAAACAGCTCTAATGTTATAACAT
ATCTTTTAACTTGGCACTGCGCCAATACTCGAATTGATATCAATCGCCGA
ATTAAGACTTCGTCTGCTAACTCTCTCTTTAGACTATGAAAAGGAGCAGA
GCTGATTGCTCACCGTCTCTCCGCTGTCCAGAAGTTTTCGTATTTTCTA

ZNF91 ko#9
>170914-066_G24_1AA4PAA028_D12.ab1 861
GGGTGACCAGTTACAGCGTCTCTTTCTTATTCAGCCCAGTGTCTGATCAC
ATCTTCTGTCACTCAGGGCCTGAGAAGGCGGGGCCTTAAACGTTATCTAA
TCAGGGACGCTGGGCTGGGAACTGTCCAATCAGGCACAAAGCTGGAGCGG
ACAGGGCGGCTTCCGGGATGTGGCTGGGCCTTTGTCTCTCGCTGCCGCCG
AACGGGGCTCGTTGAAACCGGTGGGATGCGGCTGTGGCGAGCCTCAGGCC
TTCCTTCAGTCAGCTCCACAATCCGTGCCCGGGTTCTCCTTGCCCAGCTC
GGCCTCAGTCCCCTTCAGCGATAAGATGGCGGGCTGCGGTTGACAGGGGC
GGGCTCCCGGGGCGTCCGTGTCTCCCTTCCCTGCCCAGGTGAACTGTTGT
TCCTTGGGCCTTGGGAGGCCCCTTCTTCTTGGGGGCAAGGCTTTCTTGGC
TTTCCCCCCGGCAAGGCGGGCCCGGCCGGTTCCTTTCTTTCCCCCAAGGA
AATTTTGGTTGGGCCAAGGGGGGGAAATTCCCAACCGGGGGGGGGAAGGG
GGGGGGTTTCCCGGGTTTCCCAAAGGGGGGGGGGGGGAAAAGGGAAAAAA
ATCTTCCCCCCCTTTGGGGGAAAACCCTTTTGGGGGGGGGGAAAAGGGGG
GGGGTTTTGGGGGCCCCTCTTTTGGGCGGGGGGGCCCCCTTTTTCCCCCA
TAAACTCTTTGGGGAAAAAAAAAATTATTAAAAAAAAAAAAATAAAATTT
TTTTTTTTTGGGGGGGGGGCGGGGACGGGGGGGGCGGGGGGGGGGGGGGG
GGAAAAAGGGGGGGGGGAAAAGGGGGGGGAAAAAGGGGGGGGGGGGGACC
GGGGGGGGGGG

ZNF91 ko#13
>170914-066_I02_1AA4PAA028_E1.ab1 519
GAGTGGAGGGTACAGCGTCTCTTTCTTCATTCAGCCCAGTGTCTGATCAC
ATCTTCTGTCACTCAGGGCCTGAGAAGGCGGGGCCTTAAACGTTATCTAA
TCAGGGACGCTGGGCTGGGAACTGTCCAATCAGGCACAAAGCTGGAGCGG
ACAGGGCGGCTTCCGGGATGTGGCTGGGCCTTTGTCTCTCGCTGACGGGG
CTCGTTGAAACCGGTGGGATGCGGCTGTGGCGAGCCTCAGGCCTTCCTTC

AGTCAGCTCCACAATCCGTGCCCGGGTTCTCCTTGCCCAGCTCGGCCTCA
GTCCCCTTCAGCGATAAGATGGCGGCTGCGTTGACAGGCGGGCTCCGGGC
GTCCTGTCTCTTCCCTGCCCAGTGACTGTGTCCTGGCCTGGAGCCCTCTC
TGGGCAGCTCTGCTCCCGCAGCGCCGCGTCTCTCCCAGATTGTGCAGGGA
TCACGGGAGGGTCGTCAGGGGAGAATCCTGACTGGAGGCGCTGCCTCCTG
CATCCCACGGTCCGCTTGT

ZNF91 ko#19
>170914-066_I04_1AA4PAA028_E2.ab1 513
GAGTTGCAGGTCAGCGTCTCTTTCTTATTCAGCCCAGTGTCTGATCACAT
CTTCTGTCACTCAGGGCCTGAGAAGGCGGGGCCTTAAACGTTATCTAATC
AGGGACGCTGGGCTGGGAACTGTCCAATCAGGCACAAAGCTGGAGCGGAC
AGGGCGGCTTCCGGGATGTGGCTGGGCCTTTGTCTCTCGCTGCCGCCGAA
CGGGGCTCGTTGAAACCGGTGGGATGCGGCTGTGGCGAGCCTCAGGCCTT
CCTTCAGTCAGCTCCACAATCCGTGCCCGGGTTCTCCTTGCCCAGCTCGG
CCTCAGTCCCCTTCAGCGATAAGATGGCGGCTGCGTTGACAGGCGGGCTC
CGGGCGTCCTGTCTCTTCCCTGCCCAGTGACTGTGTCCTGGCCTGGAGCC
CTCTCTGGGCAGCTCTGCTCCCGCAGCGCCGCGTCTCTCCCAGATTGTGC
AGGGATCACGGGAGGGTCGTCAGGGGAGAATCCTGACTGGAGGTTCTGGC
TCCTGAATAAAAA

ZNF611 wt#11
>180605-083_E21_1BD4PAA020_C11.ab1 662
CACCCTGCACTCGCAACAGAAA GTGTATACATTGACCTATAGCAGAAAGA
CCGGGGACGAGACCAGCCTCAGGACGACTTTAAACCCAAAAGGAAGCGAC
CCTCGGGGCTCTCACGGGGATGTCTCAATTTGCTCTGGGCAAAGGAAGAC
GCCGGAGAATTTCCAGGTCTGTGGGGACCCCACGTCCCAGGGACAGAAGC
CCTGGGAACCTAGTTAGCGCAGACTTAATACAACACAGAGCAAAACTCAC
 CGCGGCGATATGACCTACACTCCACCCGATCCGCTTCCGGGTTTGCGTTA
ATCTGCGCGCGCAGGACAAAAGCCAGGCCTGGGCGGGAAGTGGGAGGTAG
GCGGGGCCTGGGCGAGGTAGGGGCGGGGCGCGAGGCGGAGAGACCTTGCC
TTTTAGAACCGGCAGACGGCACGGCCGGTGCGGGATCTGCGCGTCTCTCA
GCCTCGCTGCCTTCAGGCTCTGCTGTCCAGAGCTCCAGACAACTCACGTA
GGGGCTTAGGAACGGGCCACAAACTTTTCGACACTCTAACCTTCCCTTCT
CGGTTGACCTACCCCCGATACATAATTAAAAACAACATAAGTCCAACCCC
AGCGTACGCACGTTTCATCGCAGCACCCCGAGGCGAAACGAAAGACCCAG
AATTCCTTATCA

ZNF611 ko#4
>180605-083_E03_1BD4PAA020_C2.ab1 731
CGACATTTACAGTGTTAGCAGGACGAAGTTCGCAGAACAAAACCTCCTCA
GAACACCGAGTTGTAGAAGGAAGGGCTTTATTCAGCTGGGAGCATTGGCA
AGCTACTGCC TTAAAATCCAAGCTCCCCGAGTGCACAATTTCTGTCCCTT
TTAAGGGCTCACAACACTAAAGATTTCAGATGAAAGGATAGTGACTGATT
GAGCAAGCAAGGGATACGTGACAGGGGCTGCATGCACCGGTGATCAGGGA
GAGAGAGAACAGGGCAGGAGTTTCACAGTGTTCTTCTATACAATGTCTGG
AATCTATGCCTAACATCAGCTTCTAAATCATAAGTTGATTTTTAACTACT
GGGTTTAGGCCAGGCGGATCCAGGCCTGGTTTCGGGCCTGGCGCGGGGCT
GCCTGTCTTTGGTTTTACTTCCTGTTGTTTTTACTGACTATAAAACAATA
TAAAACAATGTCAGAGGGTCTCTCTCTTCTCTCAACAGGGCTTTTGTATC
ATGCCTCAGACTTTCTGGCGCTCTCGCCTCCAAAACCTGAAAACAGAGAC
GCTGGGAACCGAGGGTCGCTTCCTTTTGGGTTTAAAGTCGTCCTGAGGCT

GGTCTCGTCCCCGGTCTTTCTGCTATAGGTCAATGTATACACTTTCTGTT
GCGTAGTTTTCCTGCTGAAAACCTTGTTTGACTCCCCCCCAACAGAAGAA
AGCCCCCGGCCGGGCGGGCTGGGTTCCCGCC

ZNF611 ko#16
>180605-083_E07_1BD4PAA020_C4.ab1 726
CGACATTACAGTGGTTAGCAGGGACGAAGTTCGCAAGAAAAAACCTCTCA
GAACACCGAGTTGTAG AAGGAAGGGCTTTATTCAGCTGGGAGCATTGGCA
AGCTACTGCCTTAAAATCCAAGCTCCCCGAGTGCACAATTTCTGTCCCTT
TTAAGGGCTCACAACACTAAAGATTTCAGATGAAAGGATAGTGACTGATT
GAGCAAGCAAGGGATACGTGACAGGGGCTGCATGCACCGGTGATCAGGGA
GAGAGAGAACAGGGCAGGAGTTTCACAGTGTTCTTCTATACAATGTCTGG
AATCTATGCCTAACATCAGCTTCTAAATCATAAGTTGATTTTTAACTACT
GGGTTTAGGCCAGGCGGATCCAGGCCTGGTTTCGGGCCTGGCGCGGGGCT
GCCTGTCTTTGGTTTTACTTCCTGTTGTTTTTACTGACTATAAAACAATA
TAAAACAATGTCAGAGGGTCTCTCTCTTCTCTCAACAGGGCTTTTGTATC
ATGCCTCAGACTTTCTGGCGCTCT CGCCTCCAAAACCTGAAAAAGAGACG
CTGGGACGAGGGTCGCTTCCTTTTGGGTTTAAAGTCGTCCTGAGGCTGGT
CTCGTCCCCGGTCTTTCTGCTATAGGTCAATGTATACACTTTCTGTTGCG
TAGTTTTCCTGCTGAAAACCTTGTTCGACTCCTCCCAAACGAAAAAATGA
AGCTTGAGGGGGAAAGGGGGGTGAAG