SUPPLEMENTARY INFORMATION INVENTORY

Erk1/2 Activity Promotes Chromatin Features and RNA Polymerase II Phosphorylation at Developmental Promoters in Mouse Embryonic Stem Cells

Wee-Wei Tee[1], Steven S. Shen[1,2], Ozgur Oksuz[1], Varun Narendra[1], Danny Reinberg[1,†]

This supplemental information inventory contains:

1. **Extended experimental procedures**
2. **Supplementary Figure legends (S1 – S7)**
3. **Supplementary Tables (Tables S1, S2 and S3) – attached as separate xls files.**

EXTENDED EXPERIMENTAL PROCEDURES

Mouse ES cell maintenance and differentiation

All ESCs were grown in standard ESC medium containing LIF (1000 U/ml; Chemicon). *Erk2*$^{-/-}$ ESCs were a kind gift from Dr. Sylvain Meloche. *Erk1/2* mutant ESCs were maintained in the presence of 1 μg/ml puromycin and *Erk2*-rescue ESCs in 150 μg/ml hygromycin and 1 μg/ml puromycin. For differentiation assays, ESCs were grown in the absence of LIF and supplemented with 0.1 μM all-trans retinoic acid (RA) for up to 6 days. ES to EpiSCs conversion was performed essentially as described in (ten Berge et al., 2011). MEK1/2 inhibitor (PD0325901; Stemgent) was used at a final concentration of 1 μM and GSK3 inhibitor (CHIR99021; Stemgent) at 3 μM.

Cloning, Transfections and shRNA knockdown

pLKO-based Erk1 lentiviral shRNA or control GFP shRNA (Thermo Scientific) plasmids were cotransfected with packaging plasmids (psPAX2 and pMD2) into 293 cells. Lentiviruses were harvested, concentrated, and used for ES cell infection. Stable knockdown ESCs were selected with 1 μg/ml puromycin. The Erk1 shRNA construct with the highest knockdown efficiency used was: pLKO-Erk1 shRNA #1: TRCN0000023184 (AATGTAAACATCTCTCATGGC) (Open Biosystems). For generation of *Erk2*-rescue ESCs, Erk2 cDNA was PCR-amplified from ESCs using primers containing appropriate restriction sites and epitope tags, and cloned into pCAGIG (GFP-marker) or pCAGIH

(Hygromycin-marker) expression vectors. Site directed mutagenesis was carried out to obtain the KDM construct. Constructs were electroporated into *Erk1/2* mutant ESCs using standard procedures, and clonal lines were obtained after reiterative GFP FACS or Hygromycin selection.

Site directed mutagenesis was carried out using QuikChange II XL Site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. For recombinant GST–Erk2 expression, the Erk2 coding sequence was cloned into pGEX-4T1 and expressed in Rosetta pLys competent bacteria. Purification of recombinant protein was done according to standard procedure. Full details of primers used for cloning and mutagenesis are as follows:

Erk2-rescue
Fwd: Xhol-Flag-HA-Erk2
CTCGAGCCACCATGgattacaaggatgacgacgataagggctccggatacccatacgatgtgccagattacgct
ATGGCGGCGGCGGCGGCGGCGGGC

Rev: Not1-Erk2
GCGGCCGCTTAAGATCTGTATCCTGGCTGGAATCT

Wild type GST-Erk2
Fwd BamHI-Erk2:
ATGCGGATCCATGGCGGCGGCGGCGGCGGCGGGC

Rev Not1-Erk2
GCGGCCGCTTAAGATCTGTATCCTGGCTGGAATCT

Erk2 DNA binding mutant (DBM)
Erk2-DBM-Fwd
ggatctccatcacaggaagatctgaattgtataataaatttta**cag**gct**cag**aactatttgctttctctcccgc

Erk2-DBM-Rev
gcgggagagaaagcaaatagtt**ctg**agc**ctg**taaatttattatacaattcagatcttcctgtgatggagatcc

Erk2 Kinase dead mutant (KDM)
Erk2-KDM-Fwd
caaagttcgagttgctatc**agg**aaaatcagtccttttgagc

Erk2-KDM-Rev
gctcaaaaggactgattttt**cct**gatagcaactcgaactttg

RNA analysis

Total RNA was isolated from cells using Trizol and reversed transcribed using Superscript III and oligo $(dT)_{17}$ (Invitrogen). Resulting cDNA was analyzed by Q-PCR,

performed using SYBRGREEN (Roche) and on the Stratagene Mx3005p platform.

EMSA

Each binding reaction was carried out with 1 ng of biotinylated dsDNA probe and 1 µg of Erk2 recombinant protein (wild type, DBM, KDM and GST) in 25 µl reaction buffer (50 mM Hepes, pH 7.9, 250 mM KCl, 2.5 mM EDTA, 25% glycerol, 5 µg BSA and 10 ng dI/dC). A 300-fold excess of unlabeled probe was used in the competition assays. Reactions were carried out for 1 h at room temperature and loaded onto a pre-run 6% native TGE polyacrylamide gel, and separated at 100V for approximately 80 min at 4^{o}C. The probes were transferred to nylon membranes, crosslinked and visualized using the LightShift EMSA kit (Pierce/Thermo).

Pou3f2 WT biotinylated probe:
Sense: 5' GAGAGAGAGAGAGAGAGAGAGAGTGAGAGA 3'
Antisense: 5' TCTCTCACTCTCTCTCTCTCTCTCTCTC 3'
Pou3f2 mutant cold probe:
Sense: 5' GAGTGAGTGAGTGAGTGAGTGAGTGTGAGTGA 3'
Antisense: 5' TCACTCACACTCACTCACTCACTCACTC 3'
GC-rich cold probe:
Sense: 5' CCTGCTTGGTCGTCGTGGTGCTAGCTCTTCTC 3'
Antisense: 5' GAGAAGAGCTAGCACCACGACGACCAAGCAGG 3'

Kinase assay

0.5 µg of RNAPII-CTD and 10 ng of active Erk2 (Millipore) were used in the kinase assay. Kinase reaction buffer: 50 mM Tris, pH 7.5, 10 mM MgAc, 0.1 mM ATP and 0.02 mM EDTA. Reactions were incubated for 1 h at 30^{o}C.

Immunoprecipitation

ESCs were resuspended in 10 pellet volumes of Buffer A (10 mM Tris, pH7.9, 1.5 mM $MgCl_2$, 10 mM KCl and 0.25 M sucrose), left on ice for 10 min and spun at 2000 rpm for 5 min. Cell pellets were resuspended in 5 pellet volumes of Buffer A + 0.2% NP40, incubated on ice for 15 min and spun at 3000 rpm for 5 min. The resultant nuclear pellet was washed once in MNase buffer (10 mM Tris, pH 7.9, 10 mM NaCl, 3 mM $MgCl_2$, 0.5 mM DTT, 10 mM Na butyrate), treated with 6 µl of MNase (Sigma, 200 U/µl) in the

presence of 1 mM $CaCl_2$ per ml of nuclei suspension, for 30 min at 37^{o}C. The reaction was quenched by addition of EDTA to a final concentration of 4 mM, incubated on ice for 5 min and an equal volume of high salt lysis buffer (20 mM Tris, pH 7.9, 840 mM NaCl and 20% glycerol) was added. The mixture was left to rotate at 4^{o}C for 30 min, spun at max speed and the supernatant (nuclear extract) dialyzed overnight in BC100 containing 20% glycerol. 300 µg of nuclear extract were used per IP. Appropriate antibodies were added and the mixture incubated overnight at 4^{o}C. Protein G Agarose beads were added to capture the immunocomplex, and washed 5 times in BC250 + 0.1% NP40, followed by western blot.

Triptolide treatment

Triptolide was obtained from InvivoGen (catalog # ant-tpl) and dissolved in DMSO. ESCs were treated at various concentrations for 30 min.

ChIP-PCR and ChIP seq

ChIP-westerns were first performed for each antibody to determine the optimal antibody:chromatin ratio. Chromatin-immunoprecipitated DNA was subjected to qPCR analysis and/or library construction for sequencing. For ChIP-seq, up to 30 ng of immunoprecipitated DNA was end-repaired (End-It-Repair kit; Epicenter), A-tailed (Klenow exo; NEB), and ligated to custom barcode adapters with T4 ligase (Enzymatics). After quantification, libraries were sequenced on Illumina HiSeq. We employed a custom barcoding system to sequence more than one sample per lane.

Computation analysis

Sequencing reads mapping and preprocessing

All raw sequencing reads were mapped to the mouse genome (NCBI37/mm9) by using Bowtie aligner with v2 and m1 parameters. For all samples, reads were kept at 36 base pairs (bp) with the best quality score. The mapped reads were subsequently sorted with samtools before further analysis. The RNA seq data (Marks et al., 2012) used in **Figure 5** was processed in the same way except duplicated reads were removed after sorting. All sequencing data generated have been deposited to the GEO under the Accession number SRP028688.

To generate the visualization files for IGV genome browser, each individual sorted bam file was converted to tdf file by using IGV tools count command, in which each read was extended by 250 bp that is approximately the length of sonicated DNA fragments. The normalized coverage tracks were then computed using the 'Normalize coverage data' function in the IGV genome browser and viewed.

Identification of enriched regions (peak calling)

To identify the enriched regions, mapped reads from different biological replicates were merged first with samtools before applying MACS peak calling algorithm (Zhang et al., 2008). The positive binding regions for Erk2, Jarid2, RNAPII (total), RNAPII(S5P) and H3K27me3 were called using MACS1.4.2c package with default parameter and cutoff.

To reduce the number of false positive Erk2 binding regions that might be identified by MACS, the findPeaks algorithm (Fejes et al., 2008) was also applied to the same sequencing reads independently. The Erk2 binding regions identified by MACS but not findPeaks were excluded from further analysis, such as subsequent annotation, motif discovery and other downstream analysis.

Peak overlapping, annotation and Venn diagram

In defining the overlapped peaks for Erk2 from different peak calling algorithms, the peaks called by both MACS and findPeaks have a maximum distance of no more than 100 bp. The bioconductor package ChIPpeakAnno was used for annotation and overlapping. To overlap Erk2 with H3K27me3 or Jarid2 peaks, the maximum distance between the two respective peaks was set to less than 200 bp. The H3K27me3 and Jarid2 peaks were overlapped with Erk2 high confidence peaks, respectively, by allowing maximum distance of 200 bp between peaks. The overlapped Erk2 to H3K27me3 and Erk2 to Jarid2 peaks were then annotated to generate the gene lists, Erk2-H3K27me3 and Erk2-Jarid2. To find out what genes are in common between Erk2-H3K27me3 and Erk2-Jarid2 that also fall within 3 kb distance from the TSS of Erk2 binding regions, the three way Venn diagram was created (**Figure 3D**). The lists of genes are in Supplementary Table S3.

Plotting and visualization of reads coverage distribution

The reads coverage on single base was calculated and normalized by the total number of mapped reads such that each sample contains 10 million reads. To generate the histogram data table for visualizing average reads coverage in a particular region, the normalized reads coverage was summarized and calculated in 10 bp sliding windows. A similar method was used to generate data matrix for heatmap, with the window size in either 50 or 200 bp. To discover the possible patterns of Erk2 or H3K27me3 in a particular genome region from different experimental conditions, the distance based hierarchical clustering method was applied to the specific heatmap data matrix (Heinz et al., 2010).

Scatterplot analysis

For generation of **Figure 6C**, Erk2 and RNAPII sequenced reads were mapped to the mouse mm9 reference genome with Bowtie using the parameters –v2 –m1, yielding between 20 and 35 million mapped reads per sample. Erk peaks were called using the MACS 1.4.1 package, and filtered to contain at least 10 tags, an enrichment score > 10, and a p-value < 1E-10. The HOMER package was used to map Erk2 peaks to Erk2 target genes if they fell within ±5 kb of a ref-seq annotated protein-coding gene. DEGseq was used to calculate the normalized read densities mapping within ±3 kb of all annotated TSS.

Motif analysis

The DNA sequences of Erk2 high confidence peaks along with 100 bp extensions from both ends were used for motif analysis. The motif analysis was performed by using MEME (Bailey et al., 2009) with maximum width of 40 bp, minimum width of 10 bp.

Gene group identification for similar RNAPII(S5P) reads coverage

For **Figure S6F**, to identify the control group of genes with similar enrichment of RNAPII(S5P) reads coverage as the Erk2-PRC2 genes, we first calculated the RNAPII(S5P) reads coverage around the TSS region (±600 bp) for all RNAPII(S5P) peak proximate genes and further ranked based on RNAPII(S5P) reads coverage. We then extracted the RNAPII(S5P) reads coverage for Erk2 PRC2 genes, and in parallel, randomly selected for the same number of Erk2-PRC2 unrelated genes matched for similar RNAPII(S5P) reads coverage. We calculated and compared the Ercc3 reads coverage for these two groups of genes.

RNA seq data analysis

RNA seq data were obtained from public data set (GSE23943) (Marks et al., 2012). Raw reads were mapped and processed as described above with additional PCR duplicates removal step. The mapped reads were then assigned to refseq gene model (mm9) with htseq tools (https://htseq.svn.sourceforge.net/svnroot/htseq/trunk). The genes were subsequently sorted by the reads counts and categorized into three to four different categories based on expression value (reads counts). The high expressor list contains the top 2000 genes with the highest reads counts, low expressor list contains the 2000 genes with the least reads counts and the silenced genes contains no reads counts. The overall gene list was used as the background control. These gene lists were used for Chip-seq average reads density plotting.

ChIP primers

Oct4	GGCTCTCCAGAGGATGGCTGAG
	TCGGATGCCCCATCGCA
Nestin	GGGGAGAAGGATTCCAGAAC
	AGAAGGCTTGGAAGGAAAGC
Fgf5	GCGACGTTTTCTTCGTCTTC
	ACGAAACCCTACCGGACTCT
Brachyury	CTTTGTTTCTTCCCGCTGAG
	GCAAACCTGGTCATTCCAGT
Gata4	AAGAGCGCTTGCGTCTCTA
	TTGCTAGCCTCAGATCTACGG
HoxA7	GAGAGGTGGGCAAAGAGTGG
	CCGACAACCTCATACCTATTCCTG
HoxB13	TTGCAGACTCCTGGTGTGAG
	TTGCGCCTCTTGTCCTTAGT
HoxC5	CAGTTACACGCGCTACCAGA
	GATTCGGAATGCCAAAGAAA
Dlx5	ACCAGGTCACTGGGGAGGAT
	CTCTCCTTTCCCTTGGCACA
HoxA3 region 1	TGCTTTGTCCCCCTAAGCAA
	CAACAGGGCTTCTTGCATCC
HoxA3 region 2	ACTGGGGTTCCAGACGGAAT
	GCAGCGATCTGCATTCACAG
Zic2	ATCGTGATCCTCGCTGCATT
	CCCCTGAGCTGGGAATAACC
Meis2	GGTGGGCTTCTTAGCCCCTA
	TCCATGGAGTTTCGGCAGAG
Tbx3	GGAACTGGAGAGAACCGGAAA
	TGAGCACTGGCTCCTCCTTC
HoxA11	CTTCGTCTTACCCGGGTTCC

	CCTACAGGCAGCGAGGAGAA
HoxC11	CTCGCCTTCCCAAAATTTCC
	AGCAGAAGTTGCCCAGGTTG
Msx1	GCTAAGTCAGGTGGGCATCG
	GGAGGACGACTGGCAGAAGA
Klf2	ACAGGGTGCAGGAGCCAGTA
	GCACCCTGGACCTTGTCATC
Hist1h1a	CAGTCGCTCAGGCTGCTTCT
	GCCGGCTTCTTGGTCTTCTT
Trim28	TATTTCTGTCCCGCCTCCAG
	CCTCCCTCGCTTCTGAGACA
Rex1	ACCCAGAGCCACAGTGGAAA
	AGGGGTCCAAGGAGCTGAAC
Polr2a	TAGGGAGGGAAAGGCTGGAG
	AGCCCTTTTCGTTCCTCTGC

QPCR primers

Eomes	GCGGGGAAAACAAACAAACA
	GCCAGCCCTACAACAAATGG
Otx2	CATGATGTCTTATCTAAAGCAACCG
	GTCGAGCTGTGCCCTAGTA
Zic2	TCCATGAGTCCTCCCCTCAG
	AGGTTGGAGCTGCTTTGTGG
Gbx2	GCAGTCGGTTGATTTTGGAG
	CGACATGGCTCAGATAGGAT
Oct4	AGATCACTCACATCGCCAATCA
	CGCCGGTTACAGAACCATACTC
Nanog	AGGCTTTGGAGACAGTGAGGTG
	TGGGTAAGGGTGTTCAAGCACT
Tbx3	AGGAGCGTGTCTGTCAGGTT
	GCCATTACCTCCCCAATTTT
Prdm14	GCATCCTGGTTCCCACAGAG
	CTGCAGAACACGCCAAAGTG
Nestin	AGAGGAAGAGCAGCAAGGCCATGAC
	TCCCTGACTCTGCTCCTTCTTCTTCAT
Fgf5	AAACTCCATGCAAGTGCCAAAT
	TCTCGGCCTGTCTTTTCAGTTC
Brachyury	TGCACATTACACACCACTGACG
	AGAACCAGAAGACGAGGACGTG
HoxB13	CATTCTGGAAAGCAGCGTTTG
	TGTTGGCTGCATACTCCCG
Gata4	TTCCTCTCCCAGGAACATCAAA
	GCTGCACAACTGGGCTCTACTT
Gata6	TGCAAGATTGCATCATGACAGA
	TGACCTCAGATCAGCCACGTTA
FoxA2	CCATCAGCCCCACAAAATG
	CCAAGCTGCCTGGCATG
Math1	CCAAGTGTGTCCAGCAGTGTG
	TTGCTCTGCATTGGCAGTTG

Mash1	TAACTCCCAACCACTAACAGGC
	TGAGGAAAGACATCAACCCAG
Gapdh	GCGGCACGTCAGA TCCA
	CATGGCCTTCCGTGTTCCT
Msx1	GCGCCTCACTCTACAGTGC
	CTCTGGACCCACCTAAGTCAG
Meis2	AAGTTGGGCAGCTTTCCTCA
	CCCCCAAGCTTTGAGTTCCT
Dlx5	GGAGAACTCGGCTTCCTGGT
	GGCAGGTGGGAATTGATTGA
HoxC11	CGGATGCTGAACCTGACAGA
	ATACTGCAGCCGGTCTCTGC

Antibodies

Antibodies	Catalogue and lot number	Application
H3K27me3	Abcam ab6002; lot GR77445 and GR59572	ChIP and western
H3K4me3	Abcam ab8580; lot GR86940	ChIP
H3K9me3	Abcam ab8898; lot GR25534	ChIP
Jarid2	In house	ChIP and western
RNAPII(S5P)	Abcam ab5131; lot 37362, and GR12226	ChIP, western, IP
Histone H3	Abcam ab1791; lot GR2485	ChIP
RNAPII total	Santa cruz SC-899; lot G1008	ChIP
Ercc3	Santa cruz sc-293; lot F0111	ChIP
Erk2	Millipore 05-157 clone 1B3B9; lot 1987672, JBC1938962 and DAM1827640	ChIP and western
pErk1/2	Cell signaling 4370	Western
Erk1/2	Millipore 05-1152	Western
Eed	In house	Western
pMEK1/2	Cell signaling 9121	Western
Ezh2	In house	ChIP and Western
Tubulin	Abcam	Western
HA	Covance	Western and IP
Oct4	BD Transduction 611202	Western
Jmjd3	gift from Y.Shi	ChIP
Utx	gift from K.Helin	ChIP

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure S1 (Related to Figure 2)

(A) ChIP-QPCR for H3K27me3 performed in ESCs grown in the presence and absence of MEK1/2 inhibitor, PD03. Selected developmental genes are shown, as well as *Oct4* that represents an active gene typically devoid of H3K27me3. Results are presented as a percentage of input material. Bar represents s.d of two replicates. *Right*, Western blots for pErk1/2 as a function of PD03 treatment of ESCs.

(B) Expression of pluripotency genes upon MEK1/2 inhibition. mRNA expression of respective target genes were normalized to *Gapdh*. Bar represents s.d of two replicates.

(C) Expression of developmental genes upon MEK1/2 inhibition. mRNA expression of respective target genes were normalized to *Gapdh*. Bar represents s.d of two replicates.

Supplementary Figure S2 (Related to Figure 2)

(A) Expression of pluripotency genes in *Erk1/2* mutant cells. mRNA expression of respective target genes were normalized to *Gapdh*. Bar represents s.d of two replicates.

(B) Expression of developmental genes in *Erk1/2* mutant cells. mRNA expression of respective target genes were normalized to *Gapdh*. Bar represents s.d of two replicates.

Supplementary Figure S3 (Related to Figure 2)

(A) *Left*, Heatmaps represent the distributions of H3K27me3 regions ± 10 kb relative to 29647 UCSC annotated TSS in wild type (WT) and MEK1/2 inhibited ESCs. A 200 bp bin was used in the generation of the plot. *Middle*, Western blot showing that PRC2 components are unperturbed in *Erk1/2* mutant ESCs. *Right*, ChIP-QPCR for H3K4me3 in wild type and *Erk1/2* mutant ESCs.

(B) ChIP-QPCR for H3K9me3 on select genes in wild type and *Erk1/2* mutant ESCs.

(C) Heatmaps representing three groups of H3K27me3 distribution profile in WT and *Erk1/2* mutant ESCs. To identify the differential distribution of H3K27me3 enriched regions between WT and *Erk1/2* mutant cells, H3K27me3 peaks from

both WT and mutant were overlapped, and categorized into three groups – Enriched in WT (Group A), common to both (Group B) and enriched in *Erk1/2* mutant (Group C). Only H3K27me3 peaks that are ± 3 kb from TSS are represented and used for generating the normalized reads coverage matrix crossing the entire TSS region (± 10 kb, with 200 bp bin). The data matrix is then used for hierarchical clustering analysis. Representative ChIP-seq gene tracks for select Group C genes are shown.

(D) GO term analysis of Group C genes using GREAT database.

(E) *Left*, Average Jmjd3 profile for WT and *Erk1/2 mutant* ESCs, centered on TSS. High and low expressing genes are represented. *Right*, ChIP-QPCR for Utx on active (*Polr2a*) and developmental genes in wildtype and *Erk1/2* mutant ESCs. Note the difference in Y-axis (fold enrichment) for active versus developmental genes.

(F) Average H3K27me3 profile for WT, *Erk1/2* mutant and *Erk2*-rescue ESCs, centered on TSS. Only H3K27me3 containing genes are represented. The window size for histogram is ± 3kb with 10 bp bin size.

Supplementary Figure S4 (Related to Figure 3)

(A) ChIP-Western showing specificity of Erk2 antibody for use in ChIP. A titration was performed (increasing amounts of antibody) to determine the optimal antibody to chromatin ratio.

(B) Erk2 ChIP-QPCR in wild type and *Erk1/2* mutant ESCs on select Erk2-PRC2 targets. Results are presented as a percentage of input material. Bar represents s.d of two replicates.

Supplementary Figure S5 (Related to Figure 4)

Histone H3-ChIP QPCR for select Erk2-PRC2 bound developmental genes in wild type and GSK3 inhibited (CHIR) ESCs. Results are presented as a percentage of input material. Bar represents s.d of two replicates.

Supplementary Figure S6 (Related to Figure 5)

(A) Normalized density plots of total RNAPII in wild type and *Erk1/2* mutant ESCs. Note the specific increase in RNAPII progression on gene bodies of the highly transcribed genes in *Erk1/2* mutant ESCs. No changes in RNAPII progression

were observed on Erk2-PRC2 target genes. Note also the increase in RNAPII occupancy around TSS for Erk2-PRC2 genes in the *Erk1/2* mutant ESCs.

(B) ChIP-PCR for RNAPII(S5P) in wild type and MEK1/2-inhibited ESCs. Results are presented as a percentage of input material. Bar represents s.d of two replicates.

(C) Normalized density plots of Ercc3 for different cohort of genes, based on gene expression.

(D) Western blot showing the effects of increasing concentration of Triptolide treatment on RNAPII(S5P) levels in wild type ESCs. Cells were treated for 30 min.

(E) Western blot showing comparable levels of Erk2 expression in wild type (WT) and kinase domain mutant (KDM) rescue cells.

(F) *Left*, Box plot showing two groups of genes (non-Erk2 and Erk2 only) with comparable RNAPII(S5P) levels. *Right*, Average Ercc3 profiles for these two groups of genes.

Supplementary Figure S7 (Related to Figure 6)

mRNA expression of respective developmental genes (normalized to *Gapdh*) during retinoic acid differentiation (day 0 to day 6). Bar represents s.d of two replicates.

Supplementary Tables (Tables S1, S2 and S3) – attached as separate xls files.

Supplementary Table S1 – Heatmap_Venn.xls (relate to Figure 2 and Figure S3C)

Supplementary Table S2 – Erk2 High confidence Peak Annotation.xls (relate to Figure 3)

Supplementary Table S3 – Fig 3D_Venn.xls (relate to Figure 3D)

Supplemental References

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research *37*, W202-208.

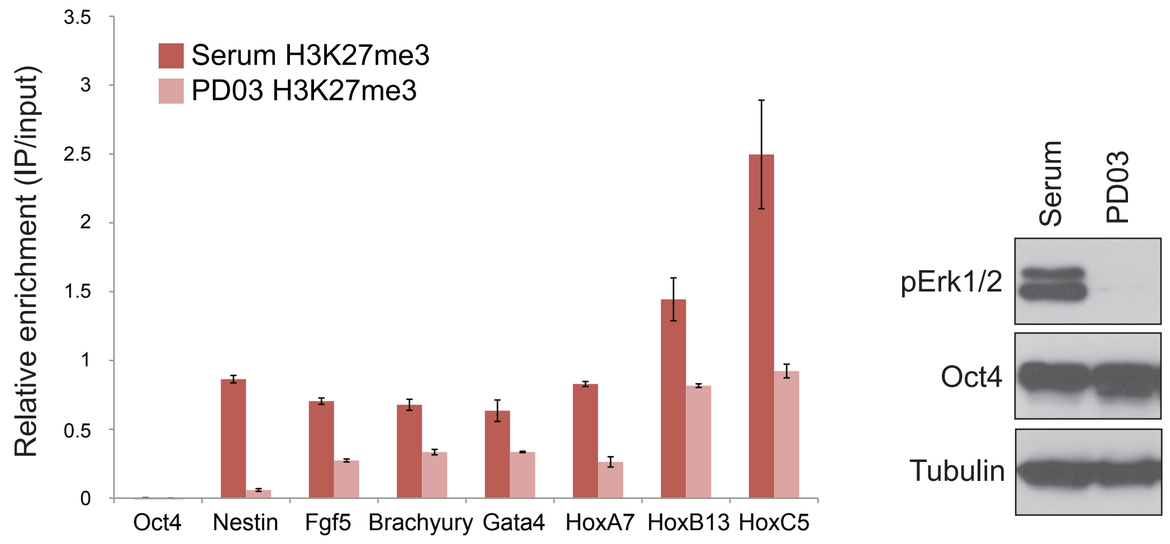
Fejes, A.P., Robertson, G., Bilenky, M., Varhol, R., Bainbridge, M., and Jones, S.J. (2008). FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics *24*, 1729-1730.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

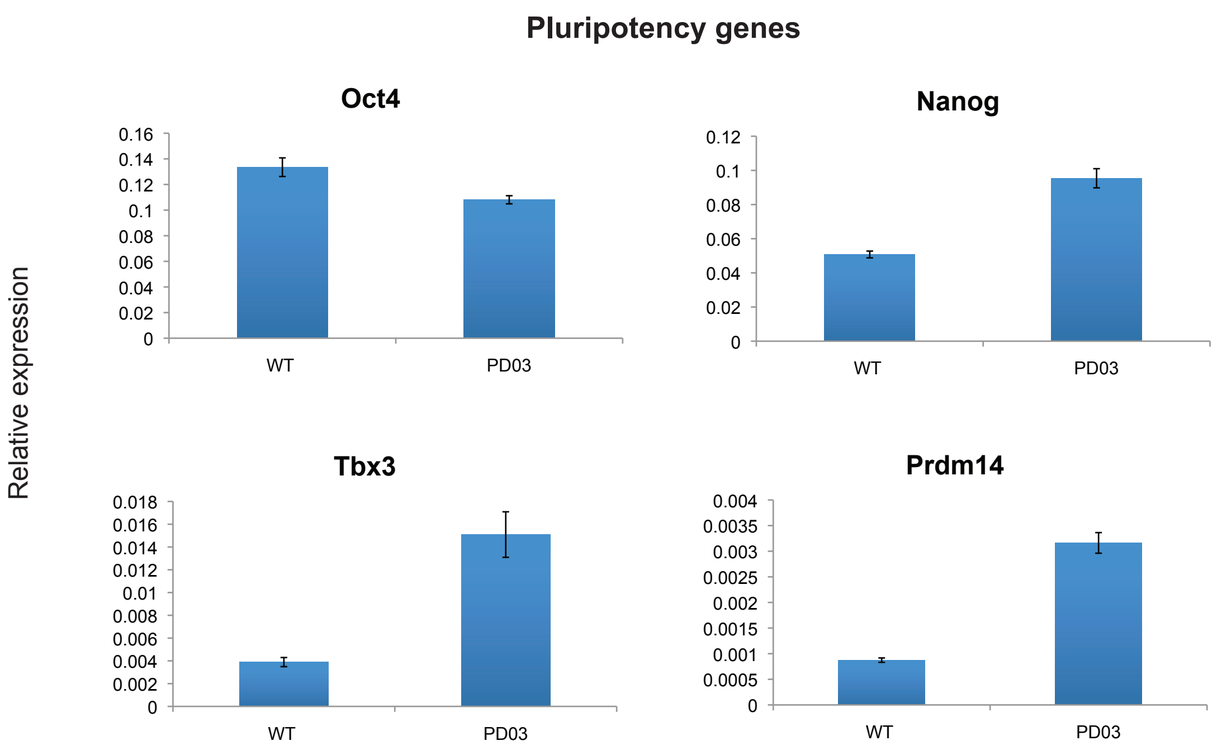
Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A.*, et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590-604.
ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nature cell biology *13*, 1070-1075.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.

SUPP FIG S1

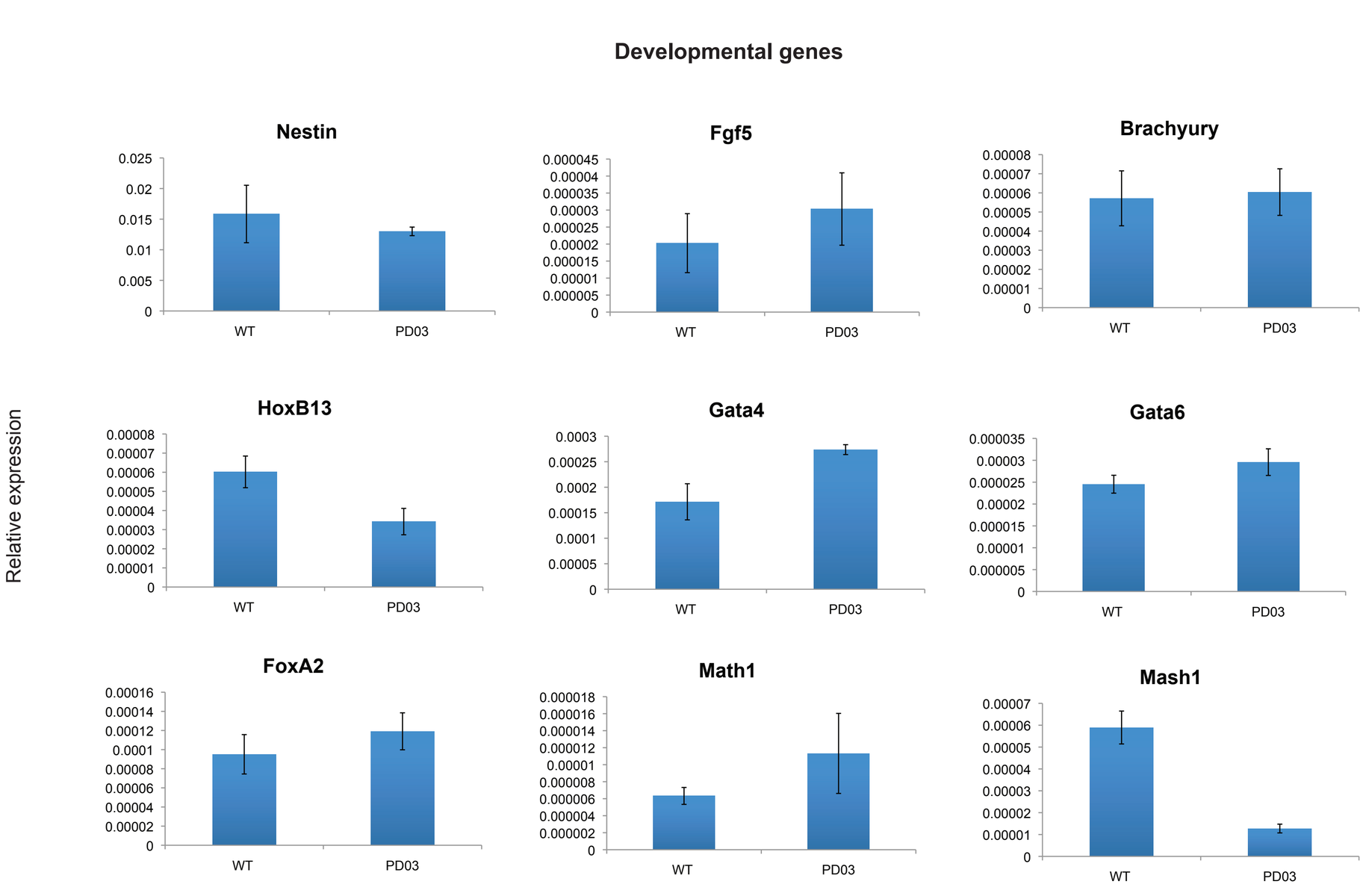
A



B

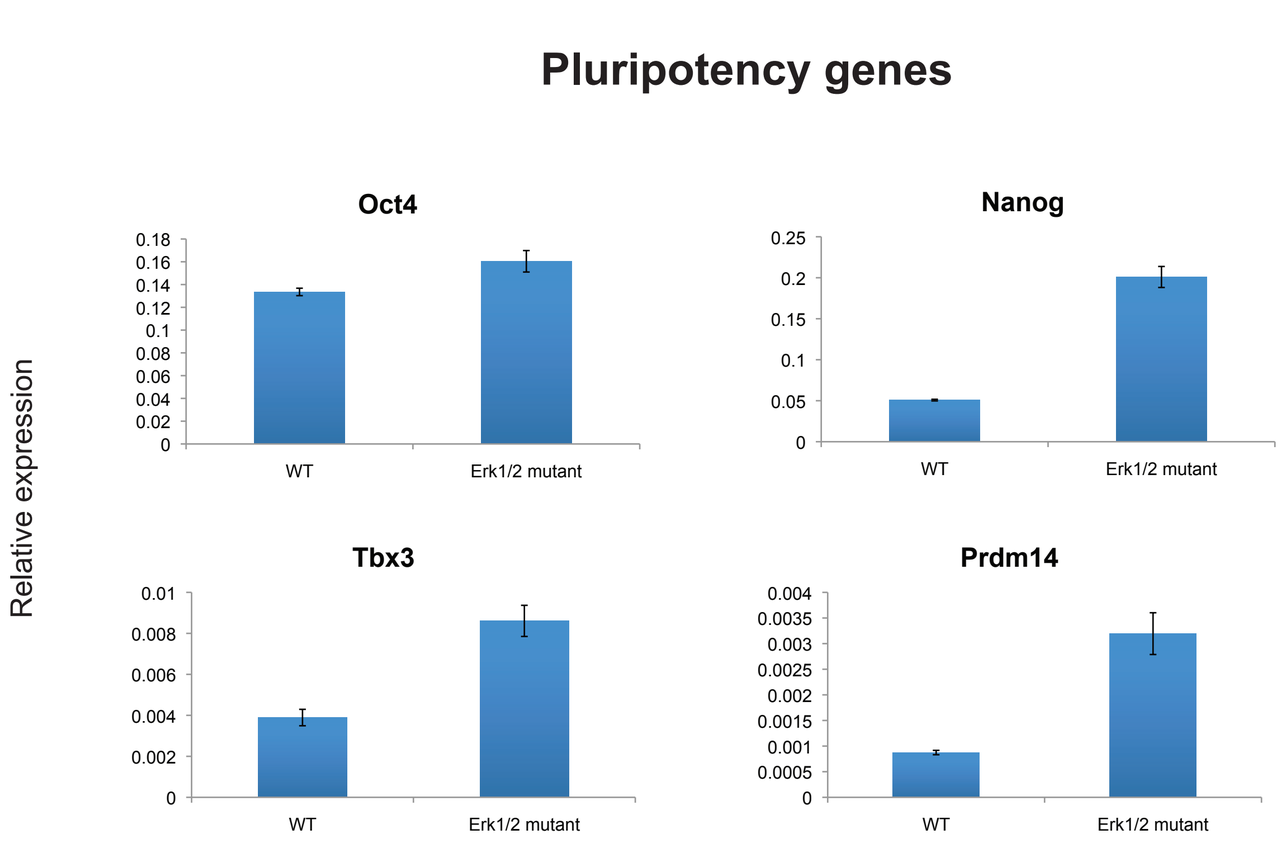


C

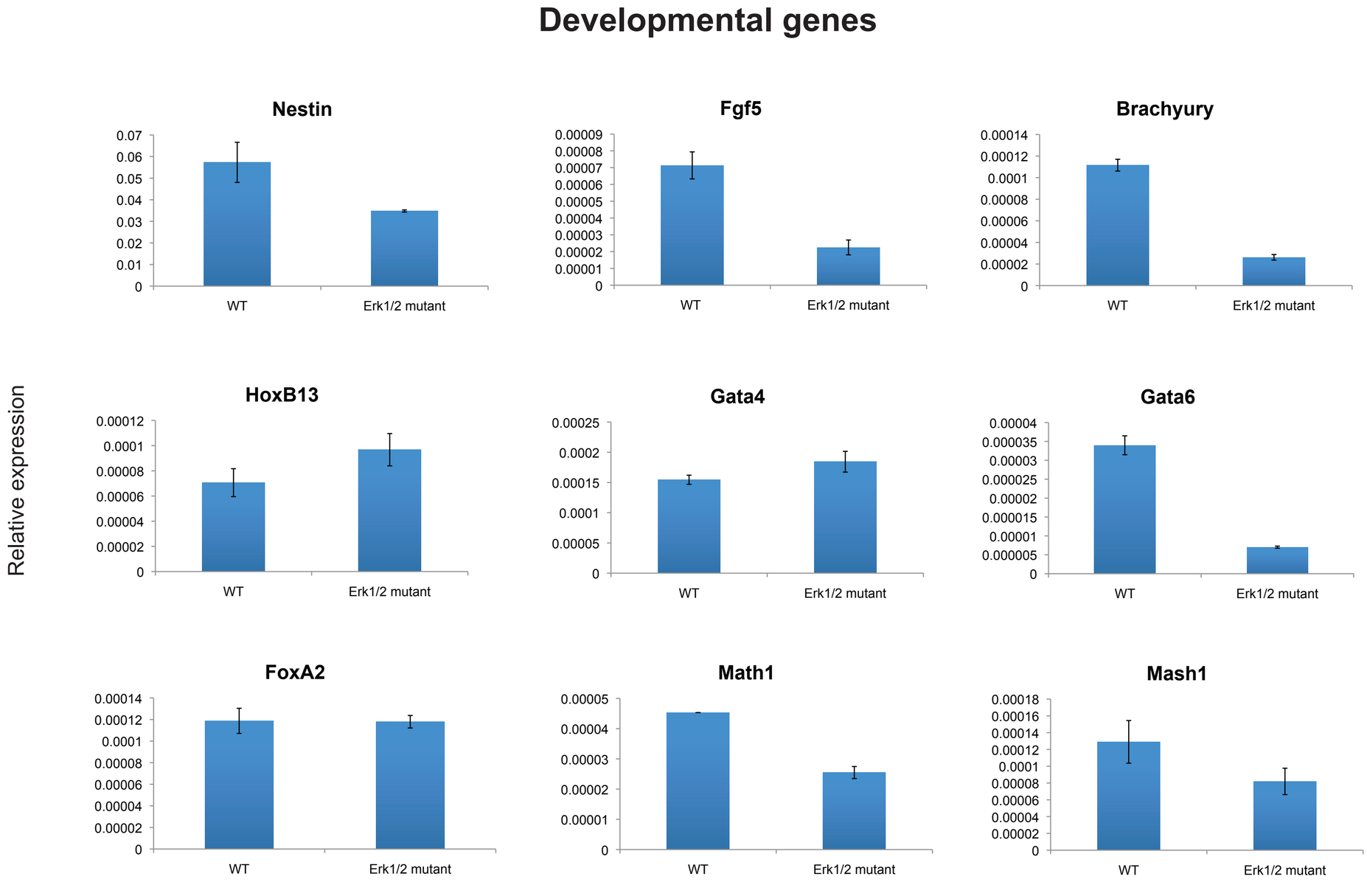


SUPP FIG S2

A

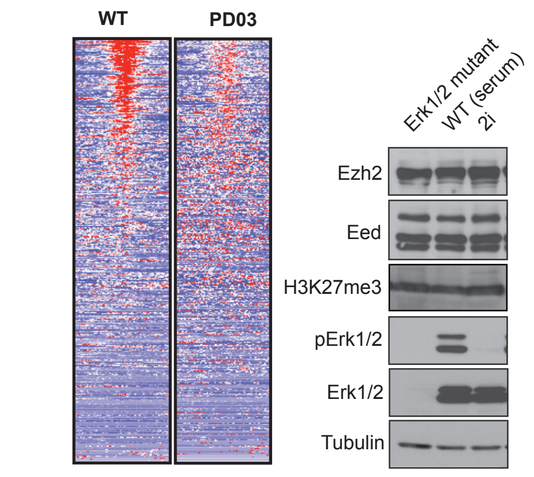


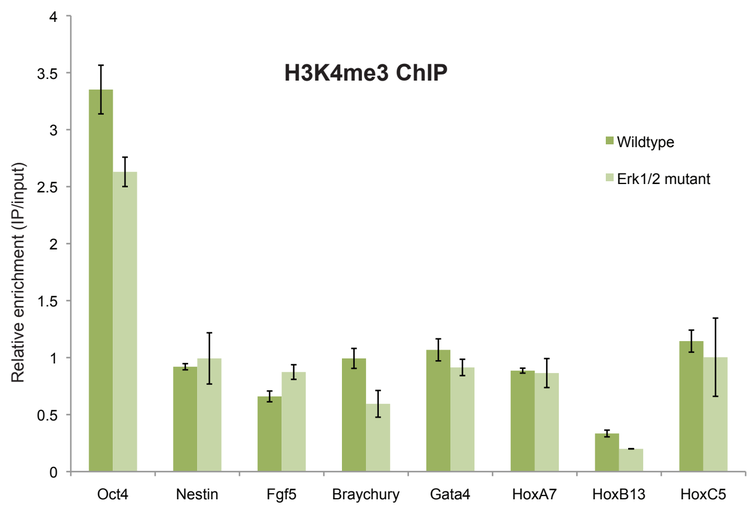
B



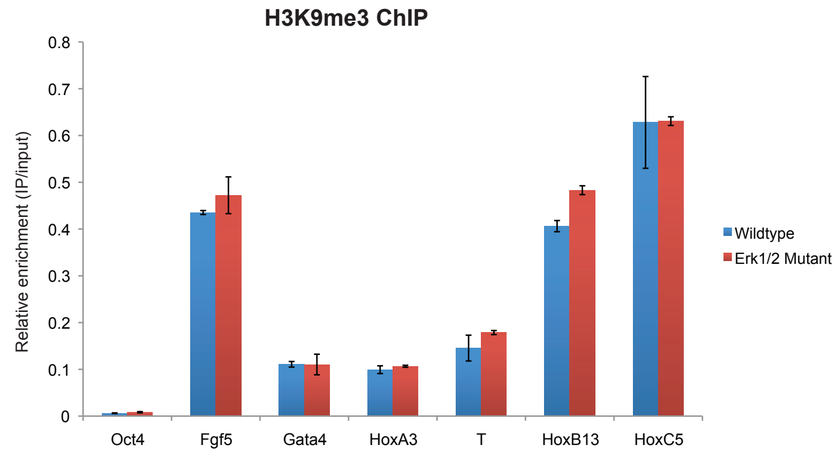
SUPP FIG S3

A

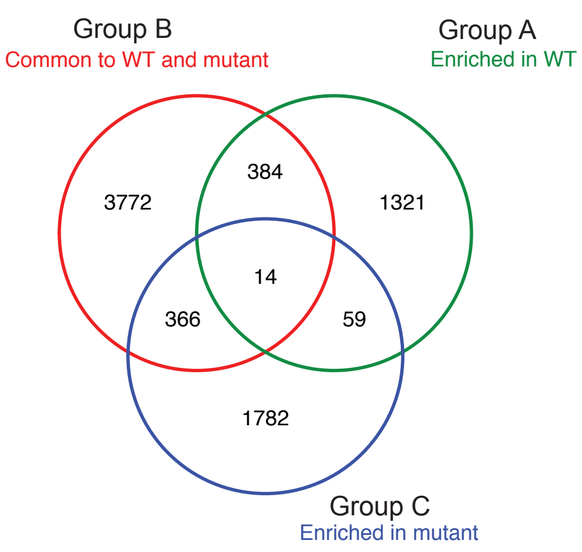


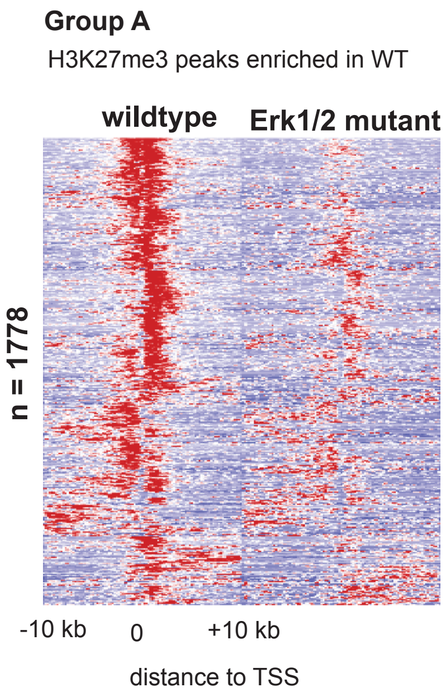


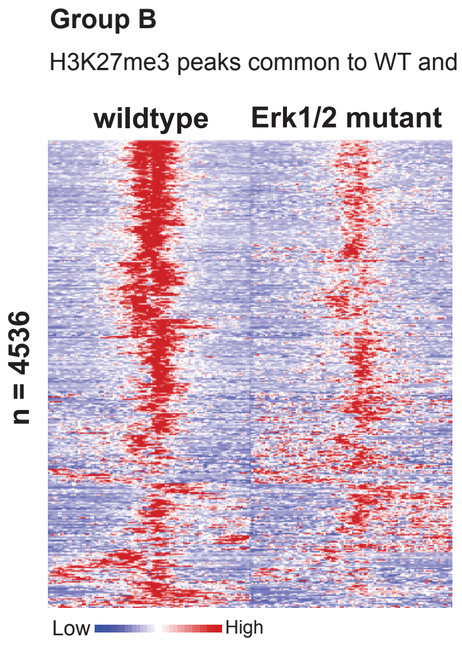
B

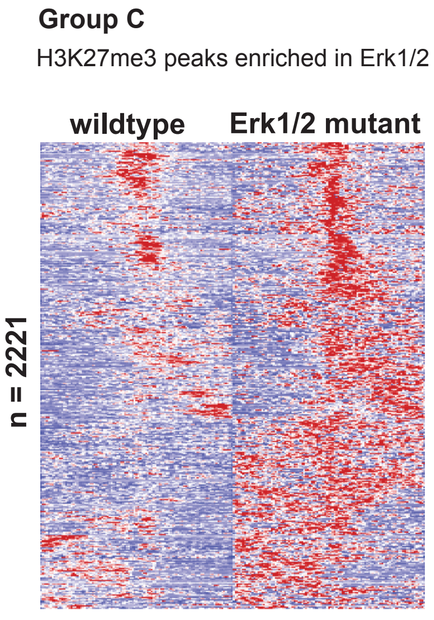


C

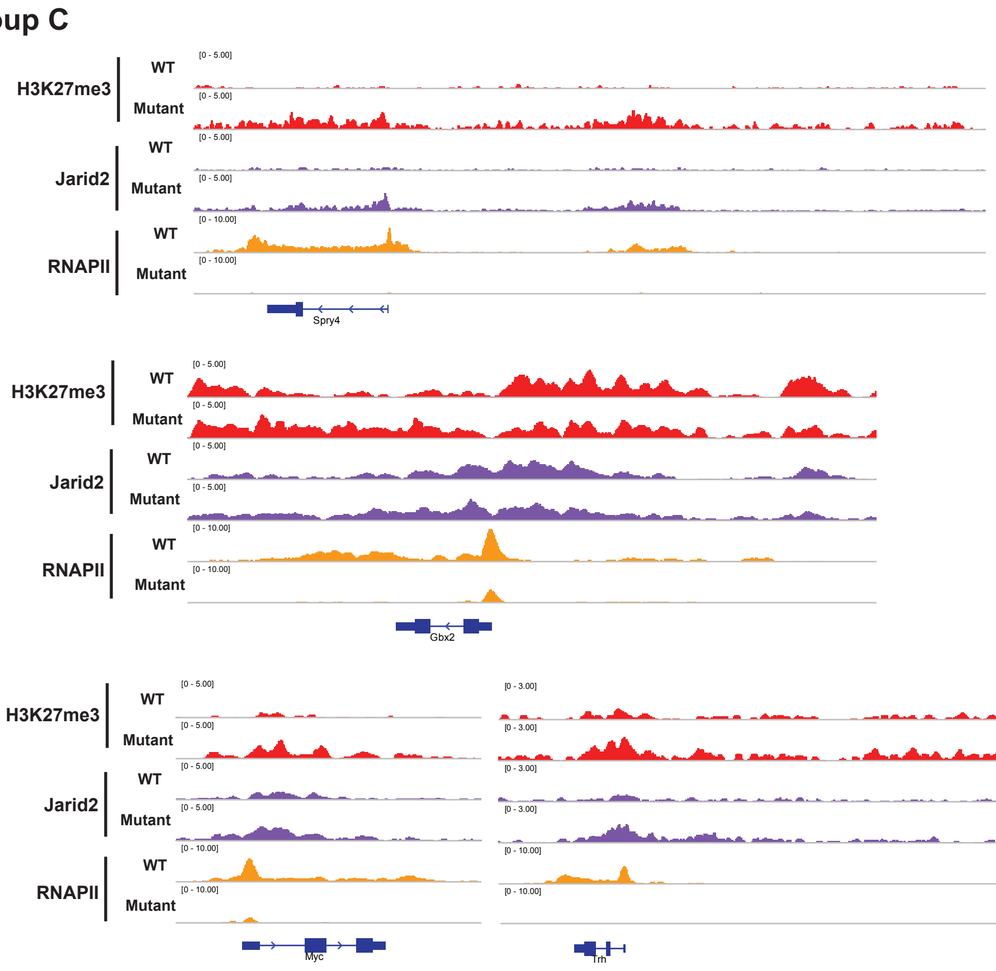


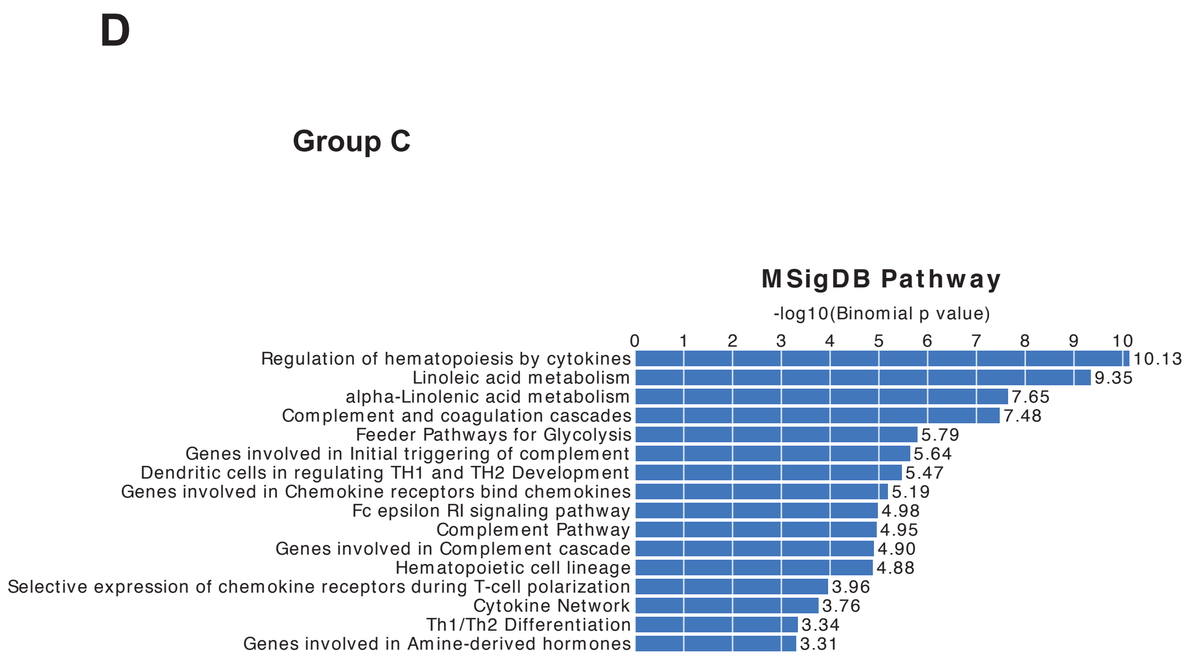




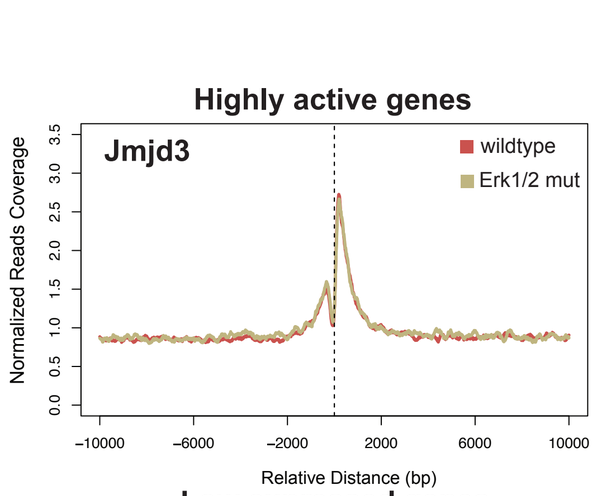


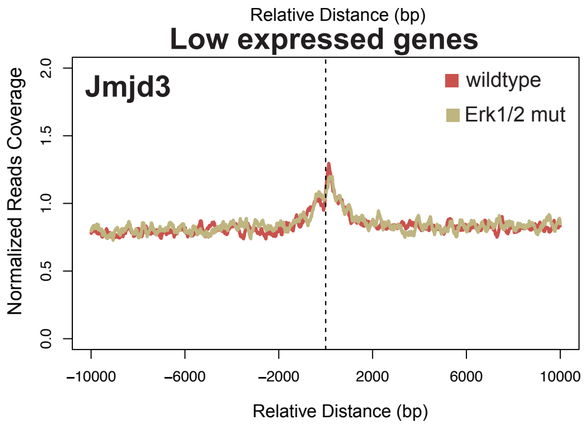
D



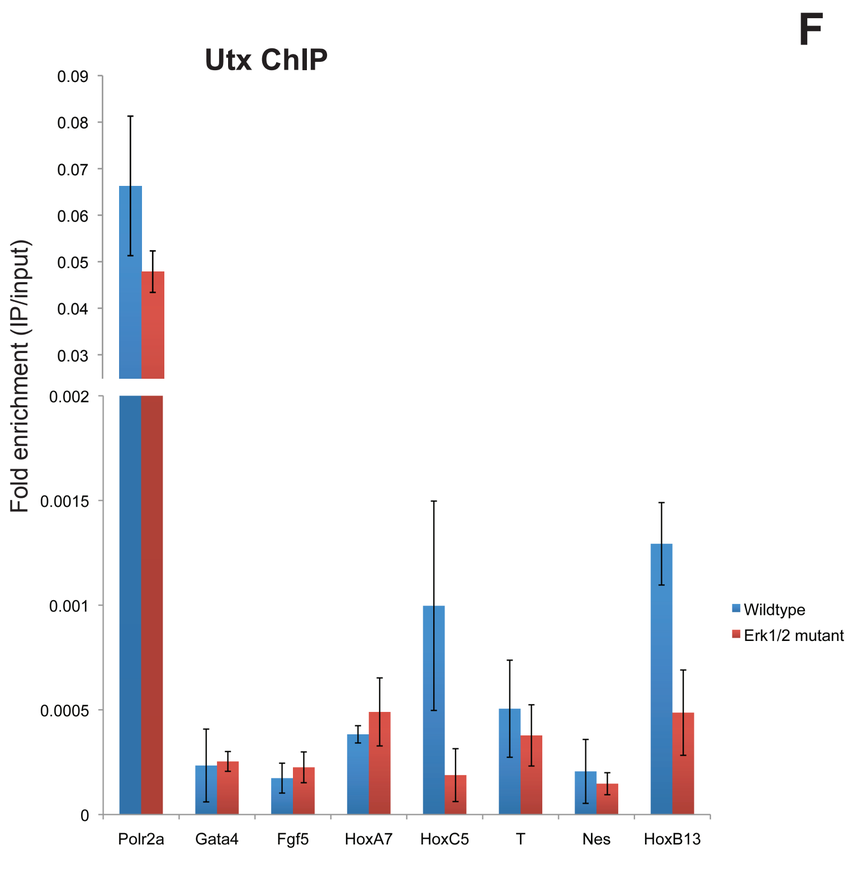


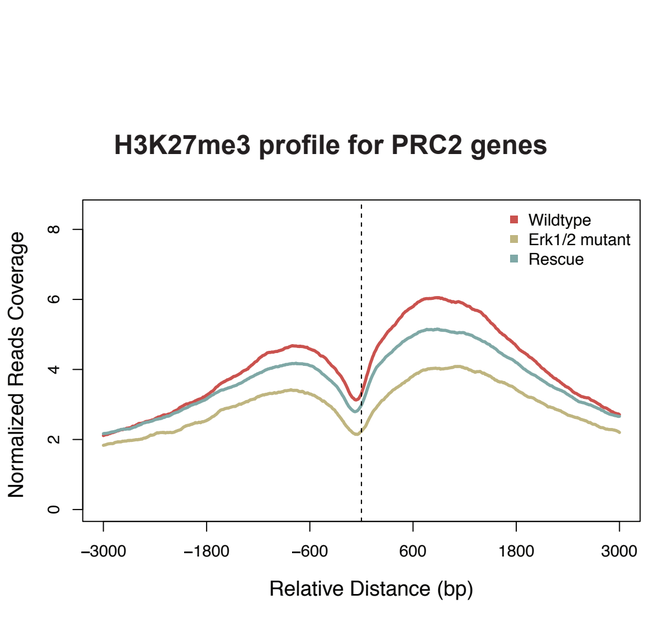
E





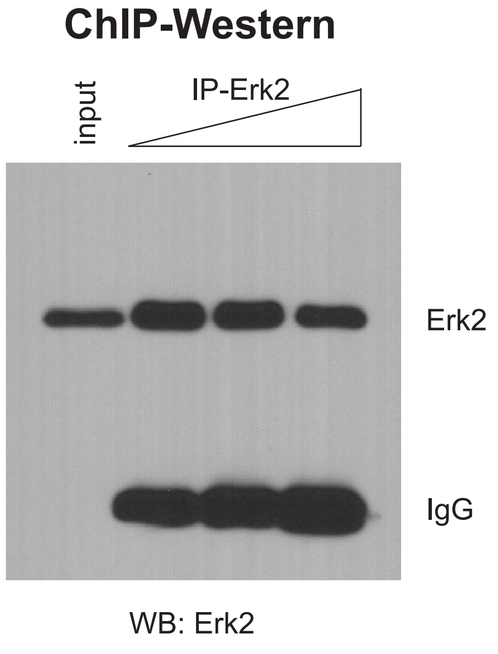
F



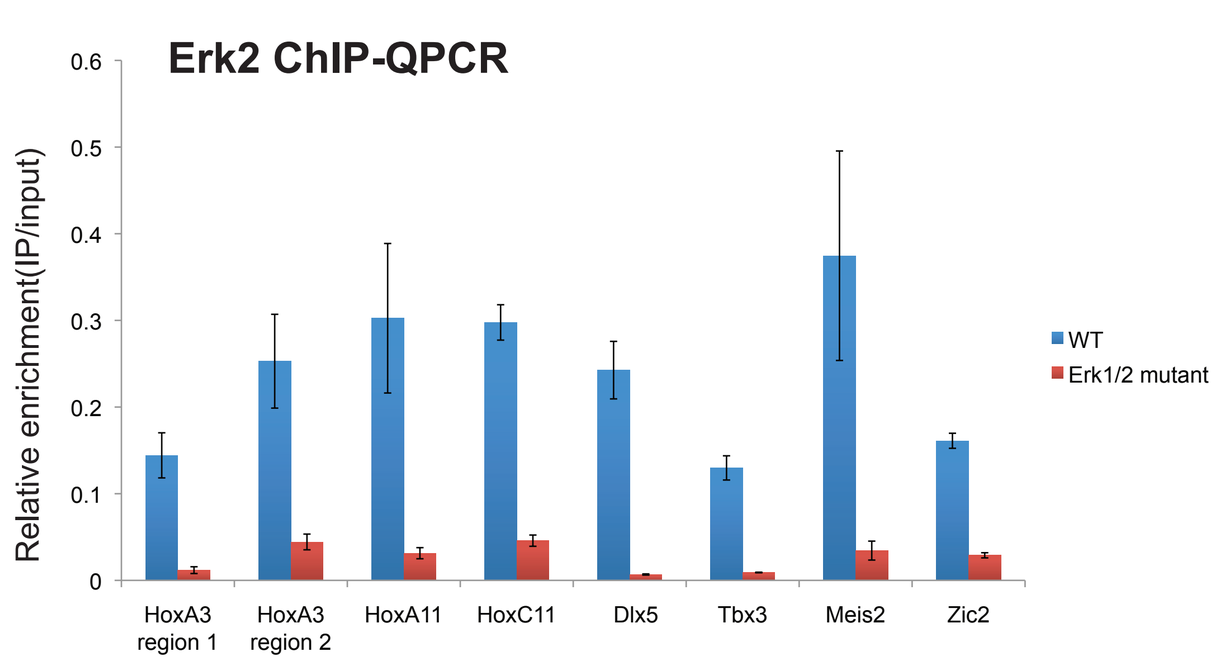


SUPP FIG S4

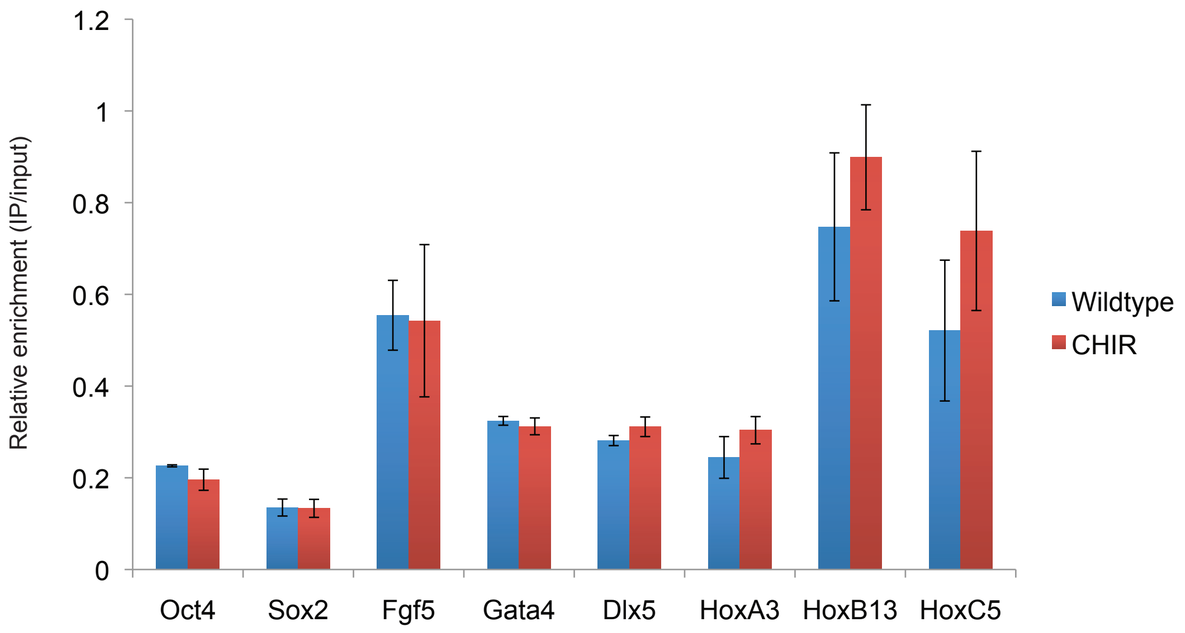
A



B

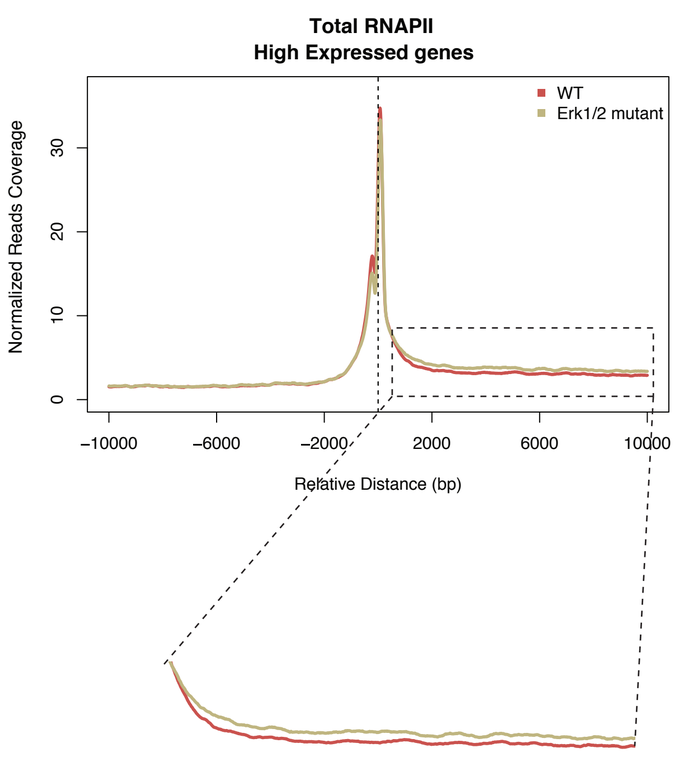


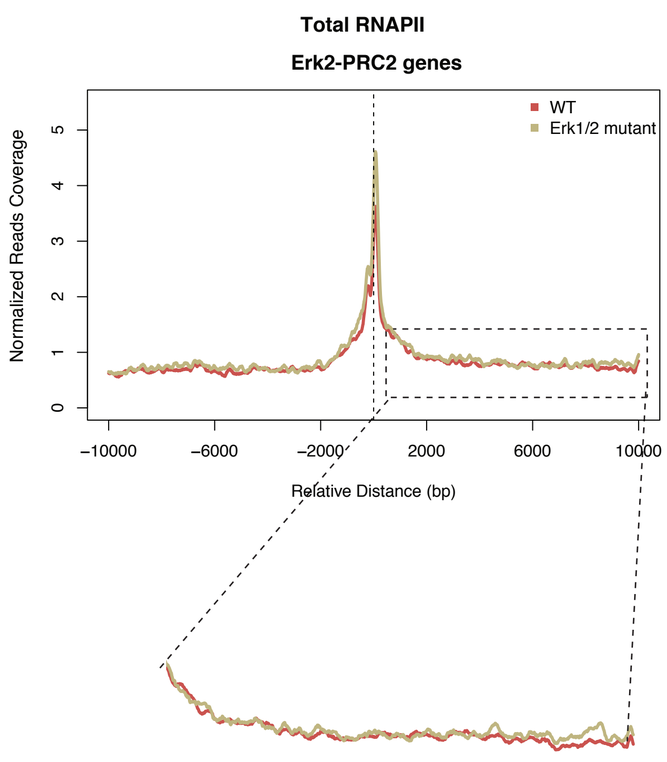
SUPP FIG S5


Relative enrichment (IP/input)
1.2
1
0.8
0.6
0.4
0.2
0
Oct4
Sox2
Fgf5
Gata4
Dlx5
HoxA3
HoxB13
HoxC5
Wildtype
CHIR

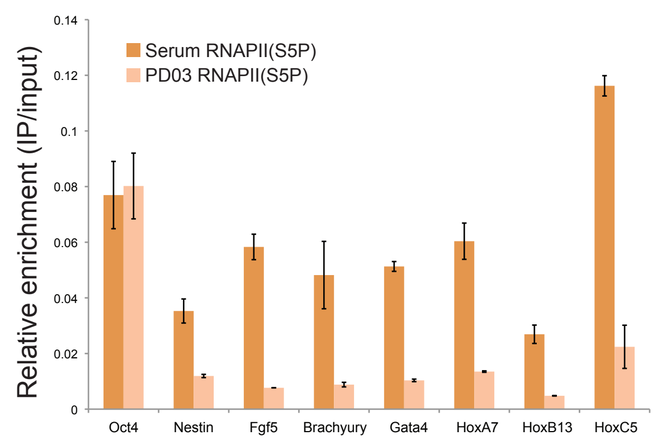
SUPP FIG S6

A

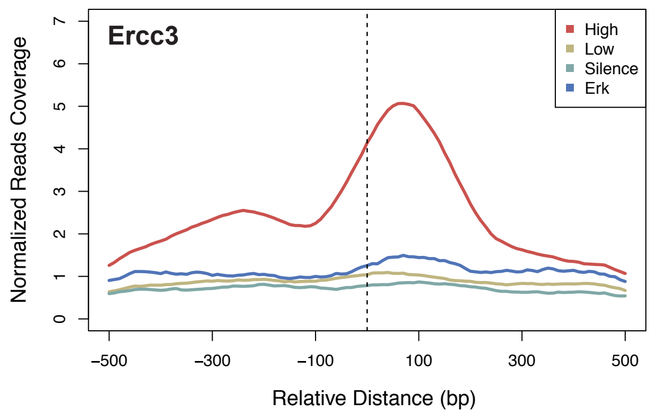




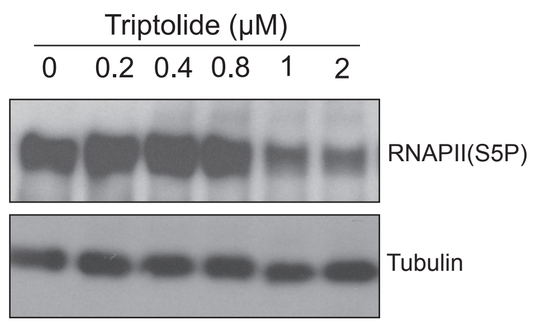
B



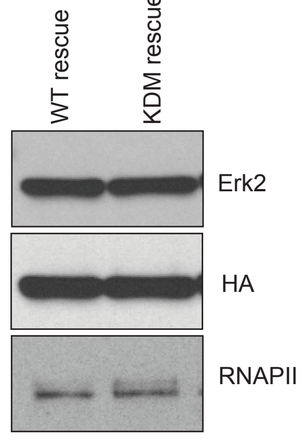
C



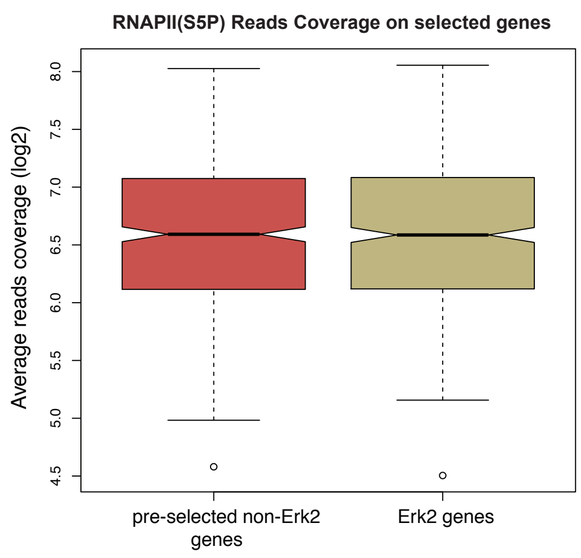
D

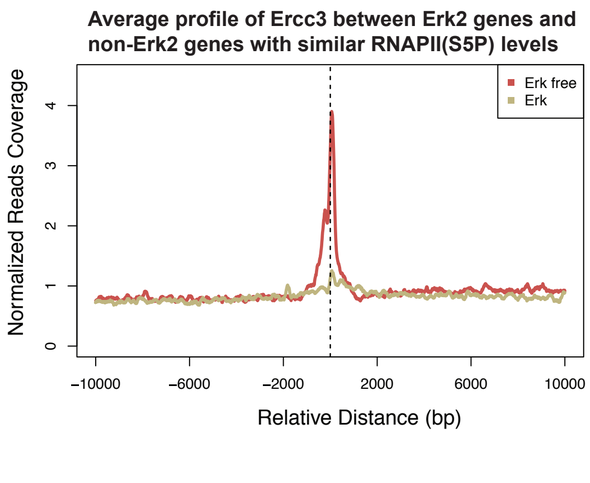


E

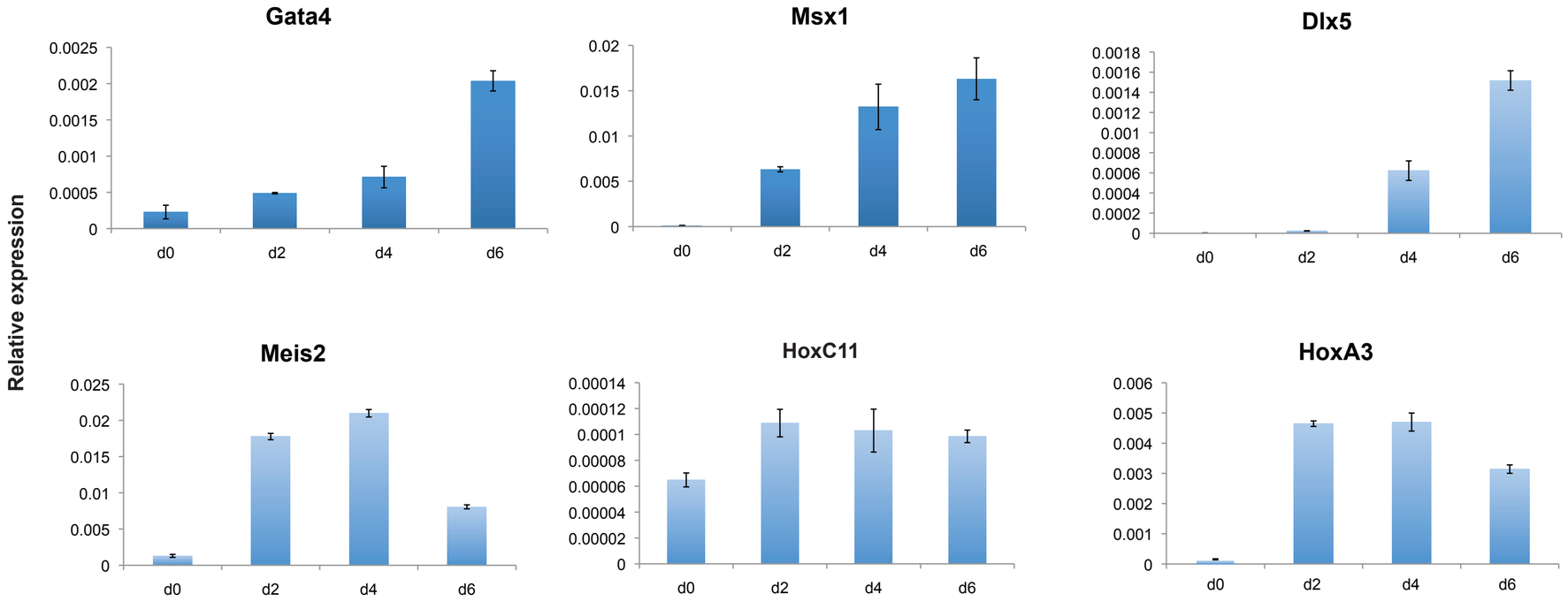


F





SUPP FIG S7


Gata4
Msx1
Dlx5
Meis2
HoxC11
HoxA3
Relative expression
d0
d2
d4
d6