**Stem Cell Reports, Volume *15***

## Supplemental Information

# UTX Regulates Human Neural Differentiation and Dendritic Morphology by Resolving Bivalent Promoters

**Qing-Yuan Tang, Shuang-Feng Zhang, Shang-Kun Dai, Cong Liu, Ying-Ying Wang, Hong-Zhen Du, Zhao-Qian Teng, and Chang-Mei Liu**

# UTX Regulates Human Neural Differentiation and Dendritic Morphology by Resolving Bivalent Promoters

Qing-Yuan Tang[1,2,3], Shuang-Feng Zhang[1,2,3], Shang-Kun Dai[1,2,3], Cong Liu[1,2,3], Ying-Ying Wang[1,2,3], Hong-Zhen Du[1,3], Zhao-Qian Teng[1,2,3,*], Chang-Mei Liu[1,2,3,*]

1. State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China.

2. Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China.

3. Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China.

*Correspondence authors: Zhao-Qian Teng, tengzq@ioz.ac.cn; Chang-Mei Liu, liuchm@ioz.ac.cn

**Running title:** UTX Regulates Human Neural Differentiation and Dendritic Morphogenesis

**Key Words:** *UTX*, Bivalent Promoters, Dendritic Morphology, Neural Differentiation

# Supplemental Figures and Legends

**A**

| Locus gene | In exon | Locus | Mismach sequence | PAM Sequence | TIDE Analysis |
|---|---|---|---|---|---|
| UTX-guideRNA1 on target | Yes | chromosome X:44873562-44873581 | GGTAGCGAGCGACACTCCGC | AGG | |
| offsite1(UTY) | Yes | chromosome Y:13479636-13479655 | GGTAGTGAGCGACACTGCGC | AGG | 0 |
| offsite2(MAGIX) | Yes | chromosome X:49165015-49165034 | TGTAGCTGGGGACACTCCGC | TGG | 0 |
| offsite3(SMPD1) | Yes | chromosome 11:6390606-6390625 | GGTACGGAGCGTCACTCCGC | CAG | 0 |
| UTX-guideRNA2 on target | Yes | chromosome X:44873638-44873657 | AGCGAGCGGCGAGAGCGAGG | AGG | |
| offsite1(SLC12A2) | Yes | chromosome 5:128084317-128084336 | AGCCAGCGGCGAGAGCGAGC | CGG | 0 |
| offsite2(ZNF691) | Yes | chromosome 1:42846639-42846658 | ATCGAGCGGCGGAGCGAGG | TGG | 0 |
| offsite3(LCLAT1) | Yes | chromosome 2:30447651-30447670 | CGCGAGCGGCGAGGGCGAGG | TAG | 0 |

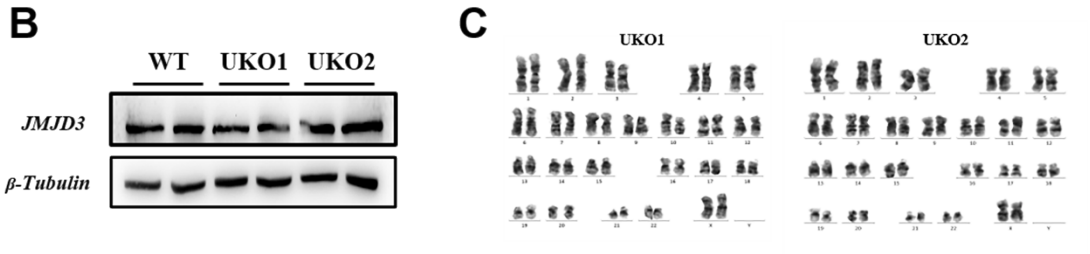


**Figure S1. UTX Deletion hESCs doesn't Induce the Off-target Effect and Abormal Karyotype, Related to Figure 2**

(A) Potential off-Targets sites of UTX-sgRNA 1/2 locus. (B) The expression of JMJD3 in UTX-KO cell lines did not change. (C) Karyotype analysis for UTX-KO hESC clones.

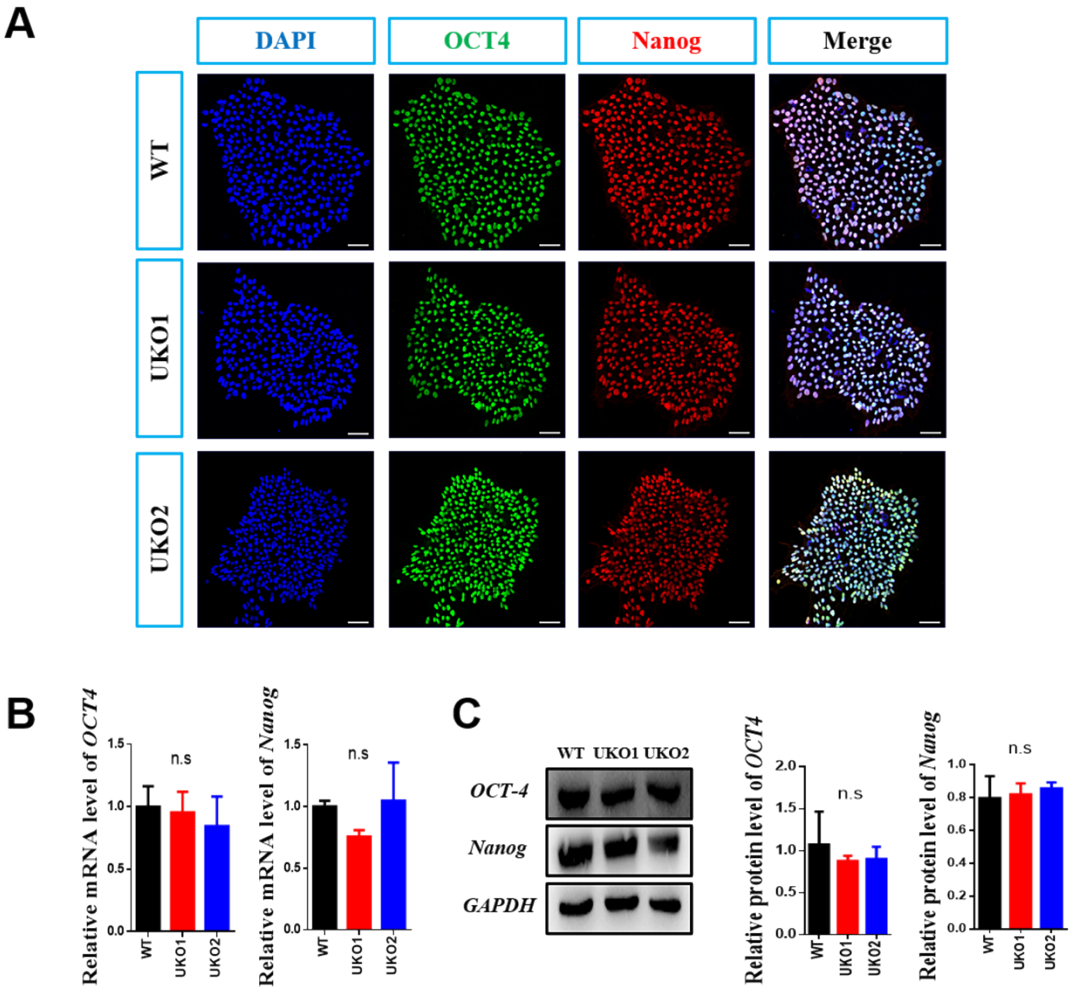


**Figure S2. The Deletion of UTX Doesn't Change the Pluripotency of hESCs, Related to Figure 2**

(A) Representative images of immunostaining results showed the expression of OCT4 and Nanog in control cells and UTX-KO clones 1/2, n=3. Scale bar: 20 μm. (B) WB analysis of control cells and UTX-KO clones 1/2 showed the expression of OCT4 and Nanog protein. (C) Real-time PCR analysis showed the expression of OCT4 and Nanog mRNA levels in control cells and UTX-KO clones. The amount of mRNA was normalized to GAPDH levels. Bars represent mean ± SEM of three experimental replicates.

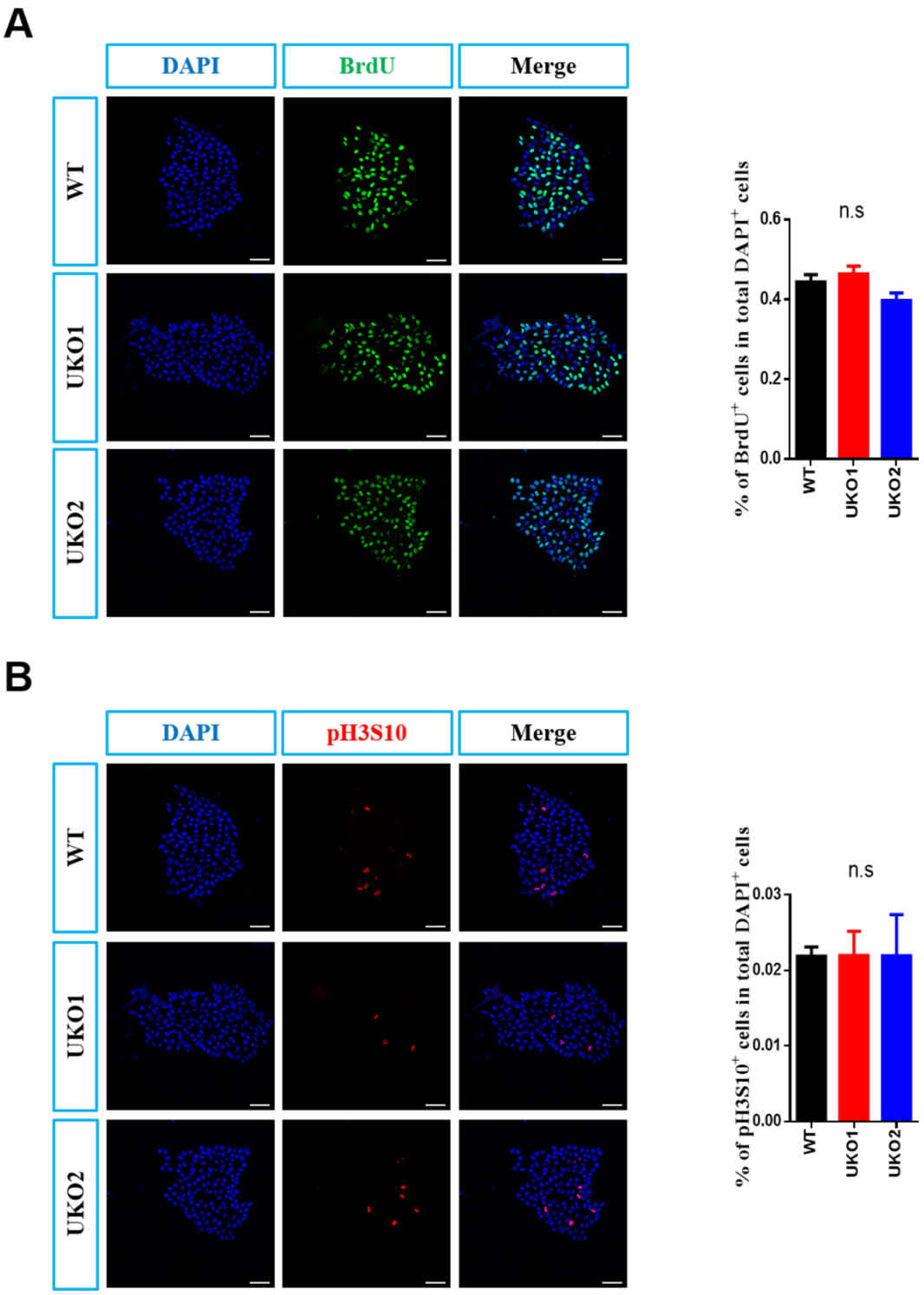


**Figure S3. The Deletion of UTX Doesn't Change the Self-renewal of hESCs, Related to Figure 2**

(A) UTX-KO and wild type hESCs immunostained positive for BrdU. Bar chart displayed percentage of BrdU cells. All error bars indicated mean ± SEM (n=3 biological replicates). Scale bar, 20 μm. (B) UTX-KO and wild type hESCs immunostained positive for pH3S10. Bar chart

displayed percentage of pH3S10 cells. All error bars indicated mean ± SEM (n=3 biological replicates). Scale bar, 20 μm.

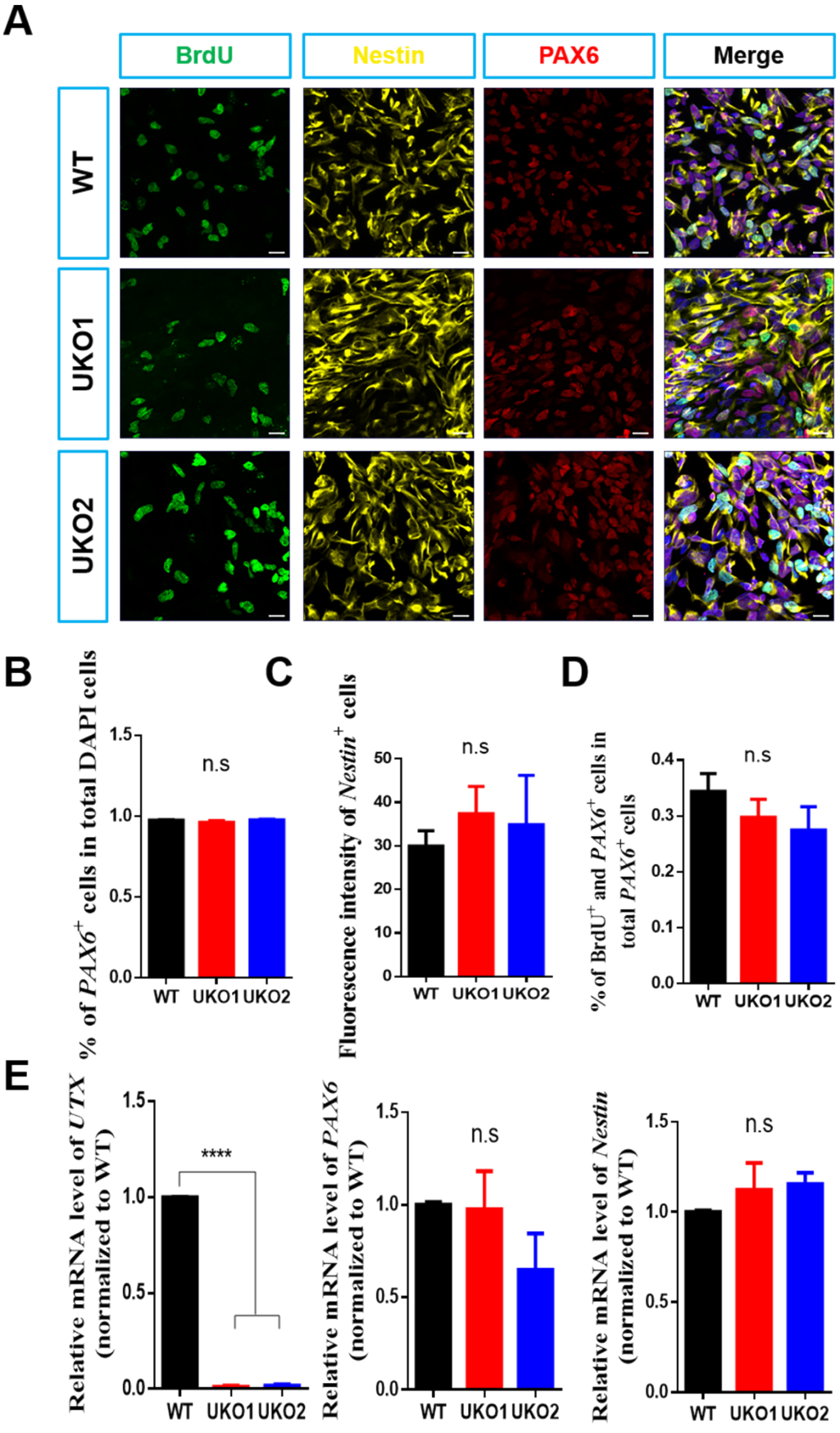


**Figure S4. Lacking of UTX Protein Doesn't Impair the Differentiation of hESCs to NSCs,**

**Related to Figure 3**

(A) 14d NSC differentiated from UTX-KO and wild type hESCs immunostained positive for BrdU Nestin and PAX6. Scale bar: 20 μm. (B). Bar chart displayed percentage of $PAX6^+$ cells. All error bars indicatedmean ± SEM (n=3 biological replicates). (C) Bar chart displayed percentage of $Nestin^+$ cells. All error bars shown indicated mean ± SEM (n=3 biological replicates). (D) Bar chart displayed percentage of $PAX6^+$ and $BrdU^+$ cells. All error bars indicated mean ± SEM (n=3 biological replicates). (E) qRT-PCR analysis for UTX/PAX6/Nestin. Results represented the averages of 4 independent differentiation experiments.

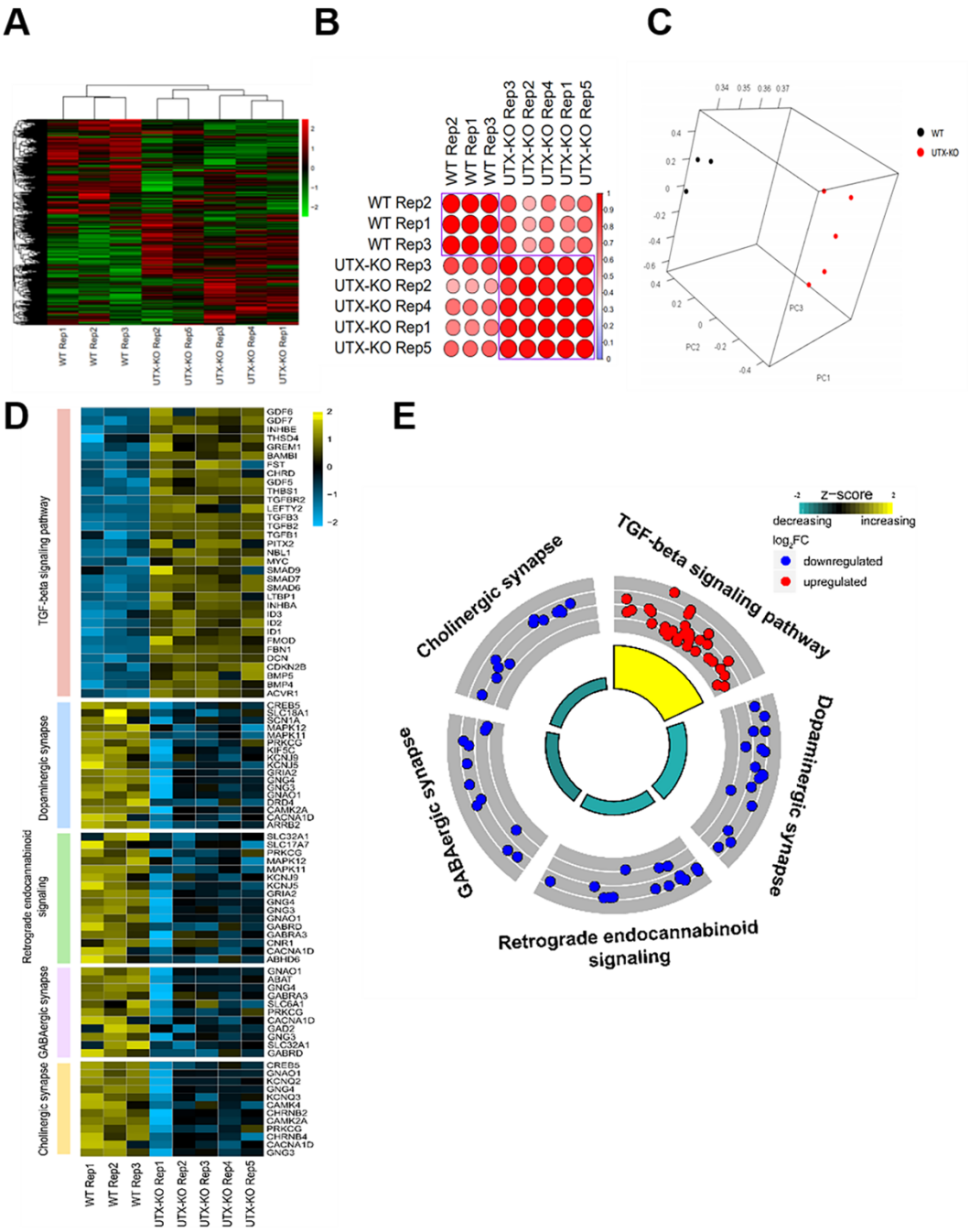


**Figure S5. Loss of UTX Results in an Abnormal Gene Transcriptional Profile in hESCs-derived Neurons, Related to Figure 5**

(A) Heatmap of RNA expression of the most differentially expressed genes between UTX-KO and WT neurons, selected by thresholds of $p < 0.05$, and $\log_{10}$ fold-change > 1 or < -1. (B) Cluster analysis of WT and UTX-KO day40 neural RNA-seq samples. (C) 3D PCA cluster analysis

of WT and UTX-KO day40 neural RNA-seq samples. (D) Heat map diagrams of differentially expressed genes in UTX-KO neurons, which were associated with KEGG pathways of TGF-beta Signaling Pathway, Dopaminergic Synapse, Retrograde Endocannabinoid Signaling, GABAergic Synapse, and Cholinergic Synapse. (E) KEGG pathway analysis revealed that the differentially expressed genes in UTX-KO neurons were enriched for multiple cellular biological processes, including TGF-beta Signaling Pathway, Dopaminergic Synapse, Retrograde Endocannabinoid Signaling, GABAergic Synapse, and Cholinergic Synapse.

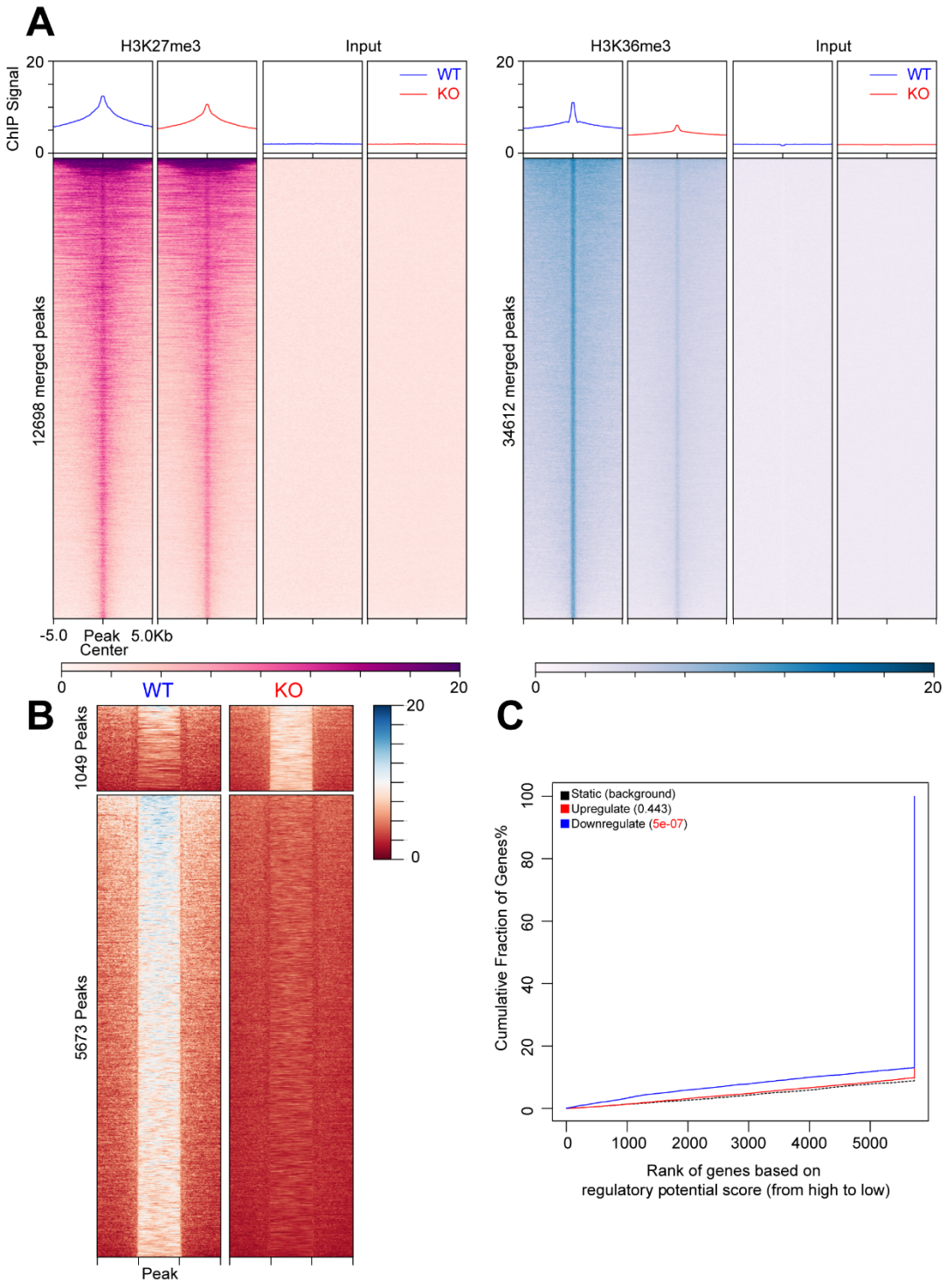


**Figure S6. The Increase of H3K27me3 is Related to the Downregulation of RNA, Related to Figure 6**

(A) Left panel, Average profiles and heatmaps of H3K27me3 peaks combing all peaks under conditions of control and UTX-KO. Right panel, Average profiles and heatmaps of H3K36me3

peaks combing all peaks under conditions of normal and UTX-KO. (B) Heatmaps of H3K36me3 peaks divided into increased enrichment (1049 peaks) and decreased enrichment (5673 peaks) group before and after UTX knockout separately. (C) BETA plot of combined computational analysis of H3K27me3 ChIP-seq and RNA-seq data (peaks with increased H3K27me3 enrichment as input, UTX-KO vs control).

# Supplemental Tables

Supplemental Table 1. Primer sequences for real-time PCR, related to Figure 2, 3 and 5

| **Gene Name** | **DNA sequence (5'-3')** |
|---|---|
| *UTX* qPCR primer-F | AGCGCAAAGGAGCCGTGGAAAA |
| *UTX* qPCR primer-R | GTCGTTCACCATTAGGACCTGC |
| Identification of *UTX* knockout qPCR primer-F | CGCCGCTTTCGGTGATGAGG |
| Identification of *UTX* knockout qPCR primer-R | CCAGTAGGGCCTTCGTCCTG |
| *OCT4* qPCR primer-F | CCTGAAGCAGAAGAGGATCACC |
| *OCT4* qPCR primer-R | AAAGCGGCAGATGGTCGTTTGG |
| *Nanog* qPCR primer-F | CTCCAACATCCTGAACCTCAGC |
| *Nanog* qPCR primer-R | CGTCACACCATTGCTATTCTTCG |
| *GAPDH* qPCR primer-F | GTCTCCTCTGACTTCAACAGCG |
| *GAPDH* qPCR primer-R | ACCACCCTGTTGCTGTAGCCAA |
| *PAX6* qPCR primer-F | CTGAGGAATCAGAGAAGACAGGC |
| *PAX6* qPCR primer-R | ATGGAGCCAGATGTGAAGGAGG |
| *Nestin* qPCR primer-F | TCAAGATGTCCCTCAGCCTGGA |
| *Nestin* qPCR primer-R | AAGCTGAGGGAAGTCTTGGAGC |
| *Tuj1* qPCR primer-F | TCAGCGTCTACTACAACGAGGC |
| *Tuj1* qPCR primer-R | GCCTGAAGAGATGTCCAAAGGC |
| *MAP2* qPCR primer-F | AGGCTGTAGCAGTCCTGAAAGG |
| *MAP2* qPCR primer-R | CTTCCTCCACTGTGACAGTCTG |
| *CREB5* qPCR primer-F | GTCAGTGAACTCCAGCATCATGG |
| *CREB5* qPCR primer-R | GTGGTGAGTCAATGCAGCCTTC |
| *CAMK2A* qPCR primer-F | GAGCCATTCTCACCACGATGCT |
| *CAMK2A* qPCR primer-R | TGGTGTTGGTGCTCTCTGAGGA |
| *PRKCG* qPCR primer-F | CCGCCTGTATTTCGTGATGGAG |
| *PRKCG* qPCR primer-R | CGATAGCGATTTCTGCCGCGTA |
| *GNG3* qPCR primer-F | GATTGAAGCCAGCTTGTGTCGG |
| *GNG3* qPCR primer-R | GGGTTCTCCGAAGTGGGCACA |
| *GNG4* qPCR primer-F | CTCCAGATTCAGCCTCCGTTTTG |
| *GNG4* qPCR primer-R | TGCCATAGGTCTGGAAGAGGTG |
| *KCNJ5* qPCR primer-F | CCTTTCTGGGAGATGTCTCAGG |
| *KCNJ5* qPCR primer-R | CCAGAGCACCTCTGTATCCATG |
| *KCNJ9* qPCR primer-F | CGCTGGTTATCAGCCACGAGAT |
| *KCNJ9* qPCR primer-R | GCTCCGAGCTTGGCATGTCATT |
| *SLC32A1* qPCR primer-F | CTGGAACGTGACCAACGCCATC |
| *SLC32A1* qPCR primer-R | TCATTCTCCTCGTACAGGCACG |
| *BMP4* qPCR primer-F | CTGGTCTTGAGTATCCTGAGCG |
| *BMP4* qPCR primer-R | TCACCTCGTTCTCAGGGATGCT |
| *BMP5* qPCR primer-F | CTCTCATCAGGACTCCTCCAGA |
| *BMP5* qPCR primer-R | GGAAGCTCACATAGAGTTCGTGC |

| | |
|---|---|
| *TGFB1* qPCR primer-F | TACCTGAACCCGTGTTGCTCTC |
| *TGFB1* qPCR primer-R | GTTGCTGAGGTATCGCCAGGAA |
| *TGFB2* qPCR primer-F | AAGAAGCGTGCTTTGGATGCGG |
| *TGFB2* qPCR primer-R | ATGCTCCAGCACAGAAGTTGGC |
| *TGFB3* qPCR primer-F | CTAAGCGGAATGAGCAGAGGATC |
| *TGFB3* qPCR primer-R | TCTCAACAGCCACTCACGCACA |
| *SMAD6* qPCR primer-F | CACTGAAACGGAGGCTACCAAC |
| *SMAD6* qPCR primer-R | CCTGGTCGTACACCGCATAGAG |
| *ACVR1* qPCR primer-F | GACGTGGAGTATGGCACTATCG |
| *ACVR1* qPCR primer-R | CACTCCAACAGTGTAATCTGGCG |
| *FEZF1* qPCR primer-F | TTCAGCCGAGGCTCTCCTAATG |
| *FEZF1* qPCR primer-R | GCCTGAAACCTTTTCCGCACAC |
| *LHX2* qPCR primer-F | ACGCCAAGGACTTGAAGCAGCT |
| *LHX2* qPCR primer-R | TTTCCTGCCGTAAGAGGTTGCG |
| *CDK5R1* qPCR primer-F | TCATCTCCGTGCTGCCTTGGAA |
| *CDK5R1* qPCR primer-R | CTCATTGTTGAGGTGCGTGATGT |
| *CNTN2* qPCR primer-F | TACGAGTGTGAGGCGGAGAACT |
| *CNTN2* qPCR primer-R | CAACGCAGGTTGGAGCCAATGT |
| *RELN* qPCR primer-F | GTCTACCTTCCACTCTCCACCA |
| *RELN* qPCR primer-R | GTCCAGCATCACAAATCCCTCG |
| *POU3F2* qPCR primer-F | GTGTTCTCGCAGACCACCATCT |
| *POU3F2* qPCR primer-R | GCTGCGATCTTGTCTATGCTCG |
| *LHX1* qPCR primer-F | GCCAAAGAGAACAGCCTTCACTC |
| *LHX1* qPCR primer-R | GGTCGTCATTCTCGTTGCTACC |
| *GLI3* qPCR primer-F | TCAGCAAGTGGCTCCTATGGTC |
| *GLI3* qPCR primer-R | GCTCTGTTGTCGGCTTAGGATC |
| *DAB1* qPCR primer-F | TACCTTTCGGCACTGCTGCTGT |
| *DAB1* qPCR primer-R | TGACCATCTGCTGTTGGACGAG |

# Supplemental Experimental Procedures

**Immunostaining of Cells in Culture**

For immunostaining cultured cells, coverslips were washed with 1×PBS for three times, then fixed with 4% PFA for 20 min at room temperature. After blocking with 2% BSA, coverslips were incubated with the primary antibodies at 4 °C overnight, and then washed by 1×PBS thrice and labeled with DAPI (Sigma-Aldrich, St. Louis, MO, USA) and the secondary antibodies for 2 hours. The primary antibodies were as follows: UTX (Cat No. GTX121246, Genetex, 1:2000), Oct-3/4 (Cat No. sc-5279, Santa Cruz, 1;1000) Nanog (Cat No. 14295-1, Proteintech, 1:1000), H3K27me3 (Cat No. 39155, Active Motif, 1:500), Purified anti-pax6 (Cat No. 901301, Biolegend, 1:1000), Nestin (Cat No. sc-23927, Santa Cruz, 1:1000), Rat anti-BrdU (Cat No. ab6326, Abcam, 1:1000), Neuron-specific type β-III tubulin (Cat No. 801202, Biolgend, 1:1000), Map2 (Cat No. 822501, Biolegend, 1:1,000) Rabbit polyclonal to S100 beta (Cat No. Ab41548, Abcam, 1:1000) . After an incubation with secondary antibodies, coverslips were mounted on glass slides. Cells were then quantified using an LSM 710 microscope equipped with a digital camera.

**RNA-Seq Analyses**

We isolated hESCs-derived neurons at day 40 in culture. Total RNAs were extracted using TRIzol reagent, and sequenced on Illumina HiSeq 2500 system. High-quality reads of RNA-seq were quantified using Salmon (v.1.1.0) with the parameter --validateMappings --gcBias" and gene expression matrix was generated by tximport (v1.14.2) (Patro et al., 2017; Soneson et al., 2015). Differential gene expression analysis was conducted using DESeq2 (v1.26.0) and increased and reduced expression was defined by log2(fold-change) > 0.585 and log2(fold-

change) < -0.585 with and P-value < 0.05 (1.5 fold change), respectively (Love et al., 2014).

**Chromatin Immunoprecipitation (ChIP)**

ChIP was performed as described previously (Liu et al., 2010). ChIP-seq libraries were sequenced generating 50-bp single reads. Raw reads data were filtered by using Trimmomatic (v.0.36) and quality-controlled using FastQC (v. 0.11.7) (Andrews, 2010; Bolger et al., 2014). High-quality reads were aligned using Bowtie 2 (v2.4.1) to the human reference genome (v33 from GENCODE) using default parameters (Langmead and Salzberg, 2012). Samtools (v.1.1.0) was then used to convert files to bam format and filter reads mapped with parameters "-F 1804 -q 30" for single-end sequencing data (Li et al., 2009). After removing PCR duplicates using Mark Duplicates function in Picard (v.2.22.0) (http://broadinstitute.github.io/picard/) and mitochondrial reads. MACS (v.2.2.7.1) was used to call peaks (-g hs -q 0.01 -f BAM --fix-bimodal --extsize 200) relative to input sample (Feng et al., 2012). MAnorm (v.1.2.0) was then used for quantitative comparison of ChIP-Seq data (Shao et al., 2012). Increased and decreased H3K27me3 or H3K36me3 enrichment were defined by M value (log2(fold-change)) > 0.585 and M value < -0.585 with and P-value < 0.05 (1.5-fold change), respectively. Peak annotation was performed using ChIPseeker (v.1.22.1) at gene level and promoter regions was defined as +/- 1000bp of TSS (Yu et al., 2015). BEDTools (v2.29.2) "multicov" function was used for read counting within genebody regions for H3K36me3 changes analysis (Quinlan and Hall, 2010). Reads per kilobase per million mapped reads (RPKM) values for ChIP-seq analysis were calculated as follows: [(read–counts) / (region–length in kb)] / (total mapped reads in Mb). DeepTools (v. 3.4.0) ''computeMatrix,'' ''plotHeatmap,'' and ''plotProfile'' functions were used to generate of heatmaps and profile plots (Ramírez et al., 2016). For genome browser

representation, data in bigwig files generated by deepTools were visualized using IGV (v. 2.4.10) (Thorvaldsdottir et al., 2013).

ChromHMM (v.1.20) was employed to analyze chromatin-state discovery and genome annotation (Ernst and Kellis, 2012). BinarizeBam function (-f 5) and LearnModel function in ChromHMM were used to convert a set of bam files of aligned reads into binarized data files and learn chromatin state models, separately.

Gene enrichment analysis was performed using clusterProfiler (Yu et al., 2012). BETA "basic" function (-k BSF --gname2 --df 1 --da 1 -c 0.001) was used for activating and repressive function prediction of different binding peaks (Wang et al., 2013). The human reference genome sequence (v33) and gene annotation (v33) were downloaded from GENCODE (https://www.gencodegenes.org/).

# Supplementary References


Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data (Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom).

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat Methods *9*, 215-216.

Feng, J.X., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nature Protocols *7*, 1728-1740.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liu, C., Teng, Z.Q., Santistevan, N.J., Szulwach, K.E., Guo, W., Jin, P., and Zhao, X. (2010). Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation. Cell Stem Cell *6*, 433-444.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Ramírez, F., Ryan, D., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.

Shao, Z., Zhang, Y., Yuan, G.C., Orkin, S.H., and Waxman, D.J. (2012). MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. Genome Biol *13*, R16.

Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res *4*, 1521.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.

Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., Tang, Q., Meyer, C.A., Zhang, Y., and Liu, X.S. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc *8*, 2502-2515.

Yu, F., Lu, Z., Chen, B., Wu, X., Dong, P., and Zheng, J. (2015). Salvianolic acid B-induced microRNA-152 inhibits liver fibrosis by attenuating DNMT1-mediated Patched1 methylation. J Cell Mol Med *19*, 2617-2632.

Yu, G.C., Wang, L.G., Han, Y.Y., and He, Q.Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omics-a Journal of Integrative Biology *16*, 284-287.