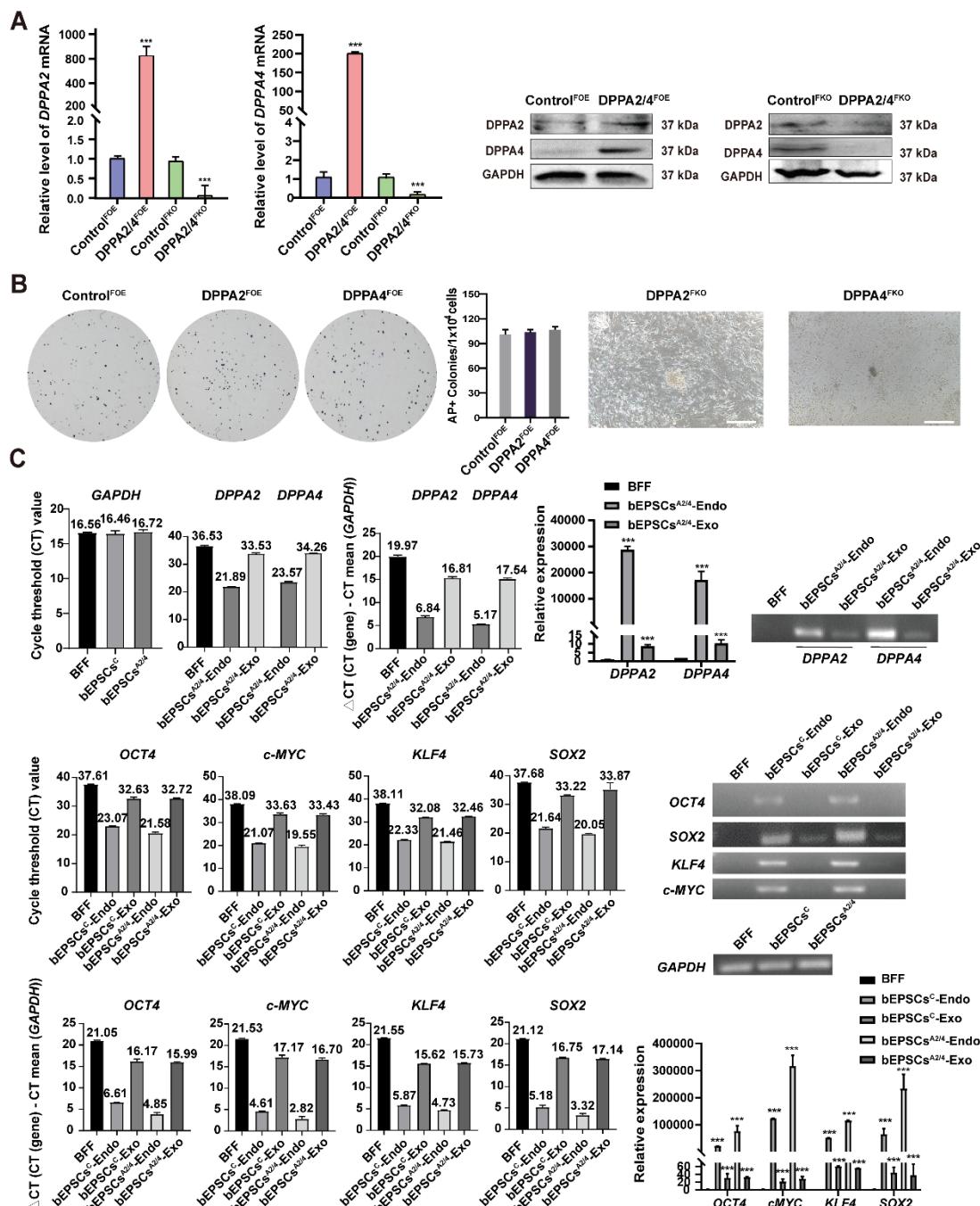


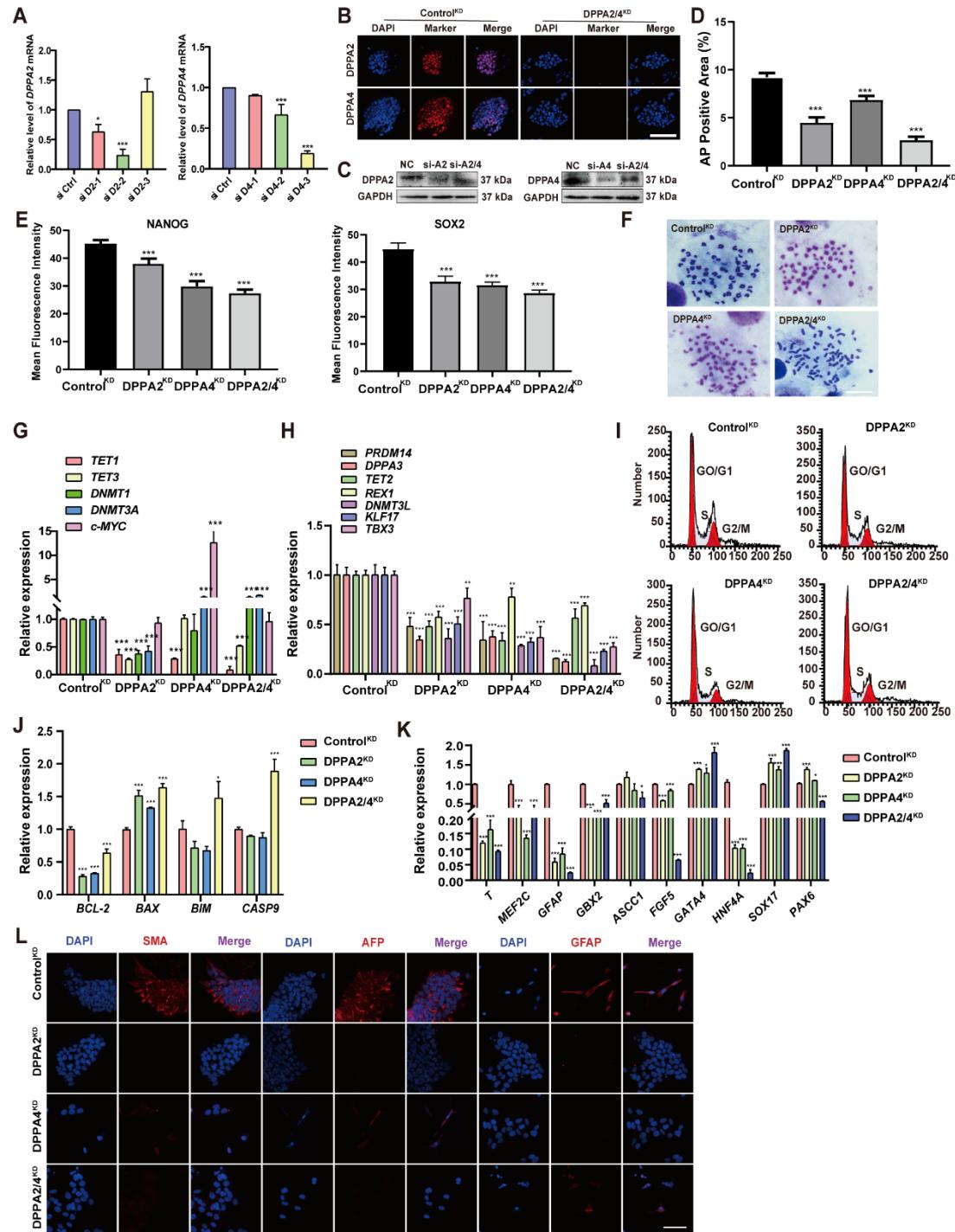
# DPPA2/4 promote the pluripotency and proliferation of bovine extended pluripotent stem cells by upregulating the PI3K/AKT/GSK3 $\beta$ /β-catenin signaling pathway

Shu Fang<sup>1</sup>, Jing Wang<sup>1</sup>, Guangbo Liu<sup>1</sup>, Burong Qu<sup>1</sup>, Jian Chunyu<sup>1</sup>, Wenqiang Xu<sup>1</sup>, Jinzhu Xiang<sup>1\*</sup>, Xueling Li<sup>1\*</sup>

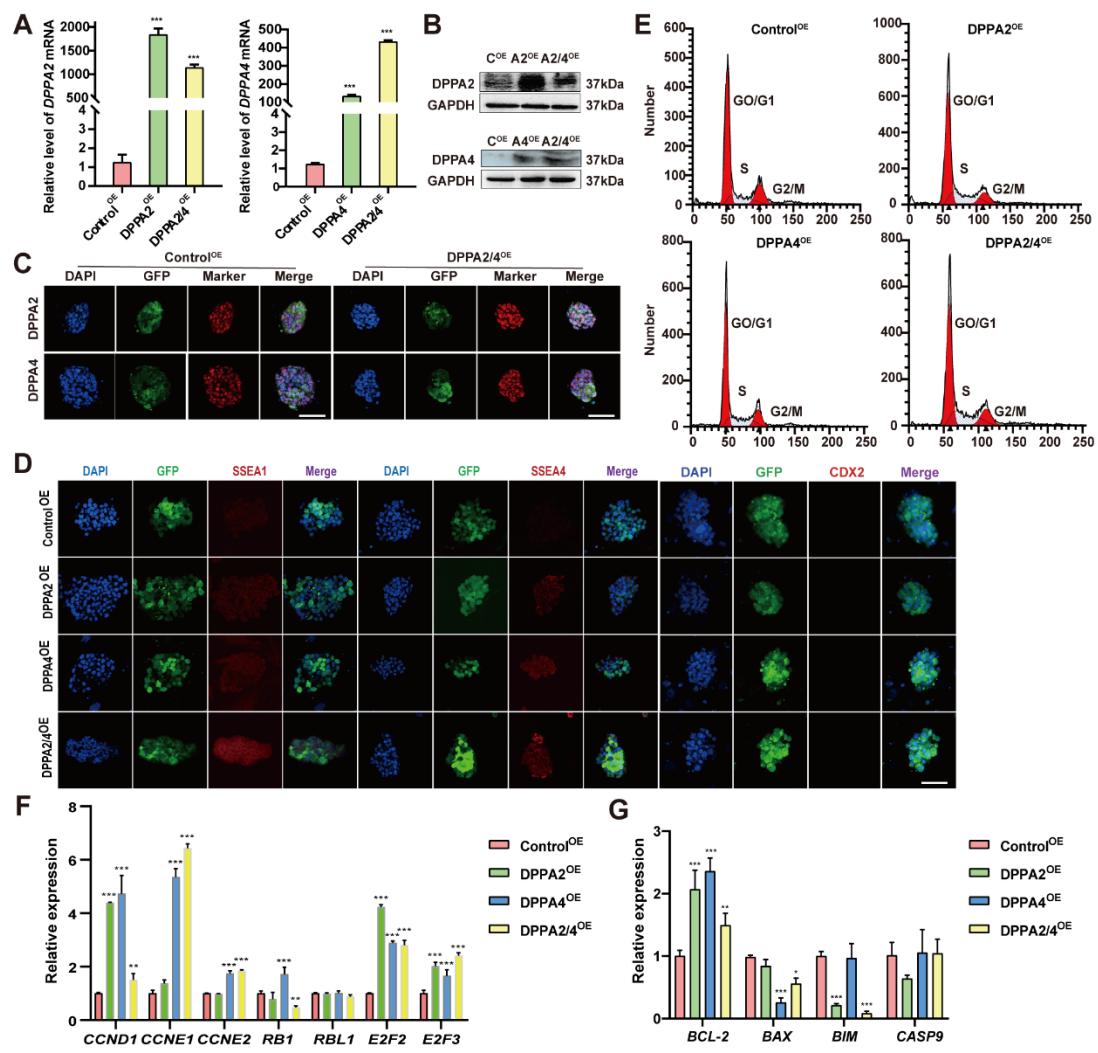


**Figure S1. Detection of the expression of endogenous and exogenous transcription**

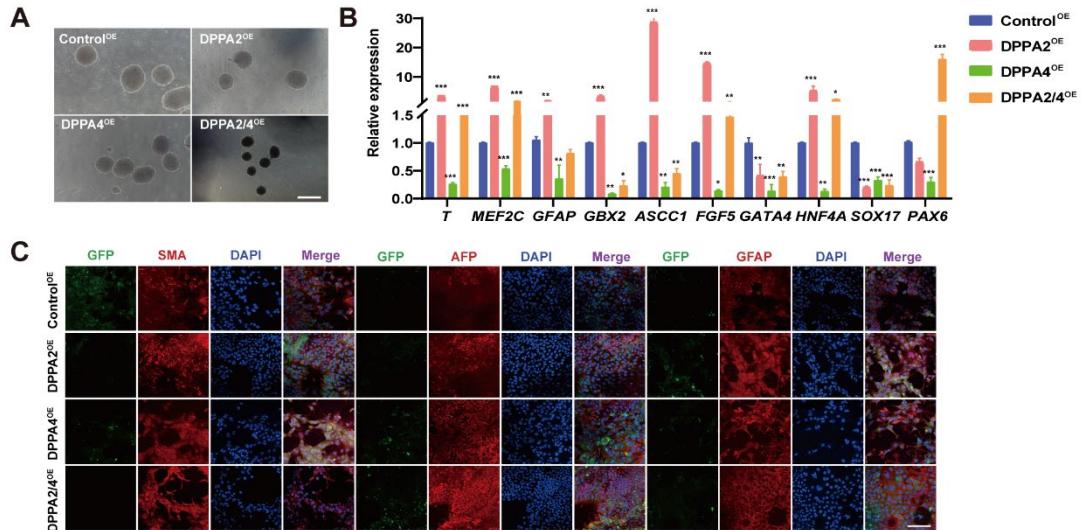
**factors after reprogramming.** (A) qRT-PCR and Western blotting analysis of the expression of DPPA2 and DPPA4 in the control, DPPA2/4<sup>OE</sup>, and DPPA2/4<sup>KO</sup> bESCs. (B) AP staining of reprogrammed BFFs from the control, DPPA2 or DPPA4-single overexpressing, and DPPA2 or DPPA4-single knockout groups. (C) qRT-PCR analysis of endogenous and exogenous genes (*OCT4*, *SOX2.c-MYC*, *KLF4*, *DPPA2* and *DPPA4*) (left), Agarose nucleic acid electrophoresis picture of endogenous and exogenous genes (*OCT4*, *SOX2.c-MYC*, *KLF4*, *DPPA2* and *DPPA4*) (right).



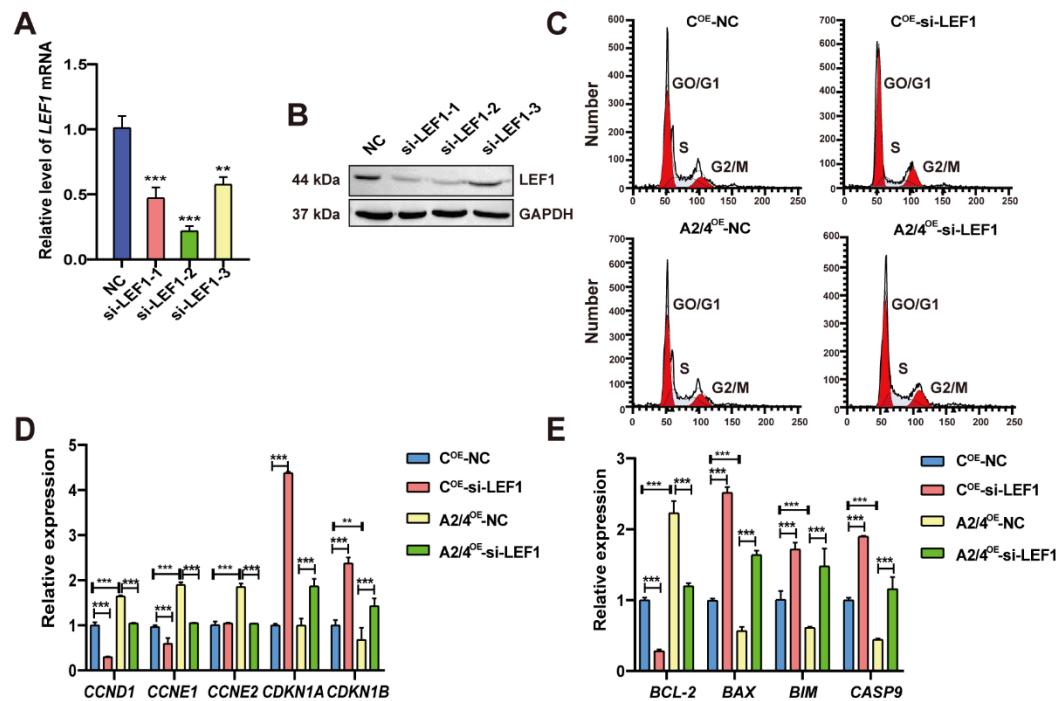
**Figure S2. DPPA2/4 knockdown affects the pluripotency and early differentiation of bEPSCs.** **(A)** qRT–PCR analysis of *DPPA2* and *DPPA4*. **(B)** Immunostaining of DPPA2 and DPPA4 in Control<sup>KD</sup> and DPPA2/4<sup>KD</sup> bEPSCs. Scale bar, 50 μm. **(C)** Western blotting analysis of DPPA2 (left) and DPPA4 (right) in Control<sup>KD</sup>, DPPA2<sup>KD</sup>, DPPA4<sup>KD</sup>, and DPPA2/4<sup>KD</sup> bEPSCs. **(D)** AP positive area (%) in Control<sup>KD</sup>, DPPA2<sup>KD</sup>, DPPA4<sup>KD</sup>, and DPPA2/4<sup>KD</sup> bEPSCs. **(E)** Mean fluorescence intensity of NANOG and SOX2 in Control<sup>KD</sup>, DPPA2<sup>KD</sup>, DPPA4<sup>KD</sup>, and DPPA2/4<sup>KD</sup> bEPSCs. **(F)** Karyotype analysis of Control<sup>KD</sup>, DPPA2<sup>KD</sup>, DPPA4<sup>KD</sup>, and DPPA2/4<sup>KD</sup> bEPSCs. Scale bar, 20 μm. **(G)** qRT–PCR analysis of the primed marker genes. **(H)** qRT–PCR analysis of the naïve marker genes. **(I)** Cell cycle analysis of Control<sup>KD</sup>, DPPA2<sup>KD</sup>, DPPA4<sup>KD</sup>, and DPPA2/4<sup>KD</sup> bEPSCs by flow cytometry. The percentages of cells in different phases are indicated. The representative histograms presented here show the distribution of cells in sequential phases (G0/G1; S; and G2/M) of the cell cycle. **(J)** qRT–PCR analysis of apoptosis-related genes. **(K)** qRT–PCR analysis of EBs at day 20. **(L)** Immunofluorescence staining for AFP, GFAP, and SMA. Scale bars, 50 μm. The data represent the means ± SDs; n = 3 independent experiments (\* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001).



**Figure S3. Overexpression of DPPA2 and DPPA4 increases the pluripotency of bEPSCs and promotes their proliferation. (A)** qRT-PCR analysis of *DPPA2* and *DPPA4*. **(B)** Western blotting analysis of DPPA2 and DPPA4 in Control<sup>OE</sup>, DPPA2<sup>OE</sup>, DPPA4<sup>OE</sup>, and DPPA2/4<sup>OE</sup> bEPSCs. **(C)** Immunostaining of DPPA2 and DPPA4 in Control<sup>OE</sup> and DPPA2/4<sup>OE</sup> bEPSCs. Scale bar, 50 μm. **(D)** Immunofluorescence staining for SSEA1, SSEA4, and CDX2. Scale bars, 50 μm. **(E)** Cell cycle analysis of Control<sup>OE</sup>, DPPA2<sup>OE</sup>, DPPA4<sup>OE</sup>, and DPPA2/4<sup>OE</sup> bEPSCs. **(F)** qRT-PCR analysis of cell cycle regulatory genes. **(G)** qRT-PCR analysis of apoptosis-related genes. The data are presented as the means ± SDs; n = 3 independent experiments (\* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001).



**Figure S4. DPPA2/4 overexpression delayed bEPSC differentiation.** (A) EB morphology of bEPSCs after DPPA2 and/or DPPA4 overexpression. Scale bars, 100  $\mu$ m. (B) qRT-PCR analysis of EBs. (C) Immunofluorescence staining for AFP, GFAP, and SMA ( $n = 3$ ). Scale bars, 100  $\mu$ m. The data are presented as the means  $\pm$  SDs;  $n = 3$  independent experiments (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ).



**Figure S5. DPPA2/4 promote bEPSC proliferation and pluripotency by upregulating LEF1 expression.** (A) qRT-PCR analysis of LEF1 gene expression. (B) Protein levels of LEF1 in bEPSCs. (C) Cell cycle analysis of C<sup>OE</sup>-NC, C<sup>OE</sup>-si-LEF1, A2/4<sup>OE</sup>-NC, and A2/4<sup>OE</sup>-si-LEF1

bEPSCs. The percentages of cells in different phases are indicated. The representative histograms presented here show the distribution of cells in sequential phases (G0/G1; S; and G2/M) of the cell cycle. **(D)** qRT-PCR analysis of cell cycle regulatory genes. **(E)** qRT-PCR analysis of apoptosis-related genes. The data are presented as the means ± SDs; n = 3 independent experiments (\* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001).

**Table S1.** The detailed information for each dataset.

| GEO accession | Samples                     | Type     |
|---------------|-----------------------------|----------|
| PRJNA693452   | 3 bovine EPSCs <sup>1</sup> | RNA-seq  |
| GSE147839     | 2 human EPSCs               | RNA-seq  |
| GSE80732      | 2 mouse EPSCs               | RNA-seq  |
| GSE110036     | 2 bovine FFs <sup>2</sup>   | RNA-seq  |
| GSE176381     | 1 human FFs                 | RNA-seq  |
| GSE127927     | 2 mouse FFs                 | RNA-seq  |
| GSE157237     | 2 human EPSCs               | ATAC-seq |
| GSE178323     | 2 human FFs                 | ATAC-seq |

<sup>1</sup> EPSCs: expanded potential stem cells. <sup>2</sup> FFs: fetal fibroblasts.

**Table S2.** The sgRNA sequences targeting DPPA2 and DPPA4

| Gene name  | Primer sequence             |
|------------|-----------------------------|
| sgDPPA2-1# | 5'- ATGTGAGTCGGGATACTTG-3'  |
| sgDPPA2-2# | 5'- AAATGGCATACTCAAATTAC-3' |
| sgDPPA2-3# | 5'- TCTGTAGTGGGTATGTGCGT-3' |
| sgDPPA4-1# | 5'- GGGATCTGCGTCGGTGTCTT-3' |
| sgDPPA4-2# | 5'- GCGATCGCAAGAGACTTGTG-3' |
| sgDPPA4-3# | 5'- GGGCTGCTCCTCCTGCACGG-3' |

**Table S3.** The primers used for the DPPA2/4 CDS amplification

| Gene name | Forward primer    | Reverse primer    |
|-----------|-------------------|-------------------|
| DPPA2     | TGGCTCTGCCTTCTC   | ACGCACATACCCACTAC |
| DPPA4     | TATCAGCCTTGACCTTT | TGCCTATAAGACCCATT |

**Table S4.** The siRNA sequences for DPPA2, DPPA4, and LEF1

| Gene name | Primer sequence |
|-----------|-----------------|
|           |                 |

|            |                               |
|------------|-------------------------------|
| siDPPA2-1# | 5'- CUCAACUUCACAAGUCAAATT-3'  |
| siDPPA2-2# | 5'- CCAAGUUCUAGCUUUAGAATT-3'  |
| siDPPA2-3# | 5'- CCUCCGAUUAACGAUGUGATT-3'  |
| siDPPA4-1# | 5'- GUUGAAAUUACAGGGAGAATT-3'  |
| siDPPA4-2# | 5'-CCAAAUGUGAUAGCAUCAATT-3'   |
| siDPPA4-3# | 5'- GUGAACCGAUCCUUCUGAATT-3'  |
| siLEF1-1#  | 5'- GCCGACAUCAAGUCUUCUTT-3'   |
| siLEF1-1#  | 5'- GGACCCUCUUACUCGAGUUTT-3'  |
| siLEF1-3#  | 5'- GCUACAUUAUGCAGCUCUAUTT-3' |

**Table S5.** The primers used in qRT-PCR analyses.

| Gene name       | Forward primer         | Reverse primer        |
|-----------------|------------------------|-----------------------|
| <i>b-DPPA2</i>  | GTCGGGATACTTTGCGGAAC   | CCATGGGATCACAGAGTCGT  |
| <i>b-DPPA4</i>  | TGAAACCTCCGAAGACACCG   | CAGCCCCTTCAGAACAGCGTT |
| <i>b-OCT4</i>   | GGTTCTCTTGAAAGGTGTT    | ACACTCGGACCACGTCTTC   |
| <i>b-SOX2</i>   | CATCCACAGCAAATGACAGC   | TTTCTGCAAAGCTCCTACCG  |
| <i>b-NANOG</i>  | TTCCCTCCTCCATGGATCTG   | ATTTGCTGGAGACTGAGGTA  |
| <i>b-CDX2</i>   | CTCCTGGACAAGGACGTGAG   | ACATGGTATCCGCCGTAGTC  |
| <i>b-DNMT3L</i> | ATGAGCAACTGGGTCTGCTT   | GGGCTCTCTTCCACACAG    |
| <i>b-DNMT3A</i> | CTGGTGCTGAAGGACTTGGGC  | CAGAAGAAGGGGCGGTACATC |
| <i>b-DNMT1</i>  | AGTGGGGACTGTGTTCTG     | TGTACGAGAGCTGCATGTCC  |
| <i>b-KLF4</i>   | TCCCACCGCTCCATTAC      | ATGAGAACTCTCGTGTAGG   |
| <i>b-KLF17</i>  | CCCTTCCCAGATGACTTTCA   | GTGGGACATCATTGGGATTC  |
| <i>b-STELLA</i> | TGCAAGTTGCCACTCAACTC   | TTCCTTGGCATAGCGAAGT   |
| <i>b-TET1</i>   | GAAACCGGAACCAATCTGTCC  | GCTTCCTCTTCTGAAAACCC  |
| <i>b-TET2</i>   | ATGCCAGTTGCAAAATTCAATT | GGGCTTCCATTCTGGAGCTT  |
| <i>b-TET3</i>   | GGAAGCGGTGTGGTACTTGT   | GCTGAGCTCTGAGCCTGTCT  |
| <i>b-T</i>      | TGCTGAAGGTGAACGTGTCT   | CACGATGTGGATTGAGGCT   |
| <i>b-MEF2C</i>  | CAGCACCAACAAGCTGTTCC   | ACCAAACGTGTTGGCTGGA   |
| <i>b-GFAP</i>   | CCTGCAGATCCGAGAAACCA   | TCCACGGTCTTCACCACAAT  |
| <i>b-GBX2</i>   | GCGAGGTGCAGGTGAAAATC   | CCCCCTGTCTTGGATTGGCA  |

|                 |                        |                          |
|-----------------|------------------------|--------------------------|
| <i>b-BMP4</i>   | GGCTGGAAGAAAAACAGAGTCC | ATGGCACTACGGAATGGCTC     |
| <i>b-ASCC1</i>  | TTCTGGATCCTTGGCCTGTAT  | CTCCACGAGGCCTTGGTAAT     |
| <i>b-FGF5</i>   | GTACGTGCCCTGAACAAGA    | GTGGGTAGAGACGTGCTGAG     |
| <i>b-GATA4</i>  | GAGATCGCCCCATCAAAAC    | TAAGGCCAGGCTGTTCCAAG     |
| <i>b-HNF4A</i>  | CGGAATCAACGGCGACATT    | AAGGCTGGGATGTACTTGGC     |
| <i>b-SOX17</i>  | CCCCTCGGGGACATGAAGAT   | TCAGTGCCTTCCACGACTTG     |
| <i>b-PAX6</i>   | GAATTCTGCAGGTGTCCAACG  | GTCTGATGGAGGCCAGTCTCG    |
| <i>b-LEF1</i>   | TTCTAGGCAGAAGGTGGCAT   | GCAGCTGTCATTCTGGACC      |
| <i>b-MYC</i>    | CCCATCAGCACAAATTACGCA  | TGTCCGCCTTGTCAATTCT      |
| <i>b-PRDM14</i> | CGGAGACAATTCCCTGATGT   | CACGGAATGTCCAGAAACT      |
| <i>b-REXI</i>   | GGAAGAGGACCCACTCCTC    | ACTTGGCCTCTAGTGCATC      |
| <i>b-DPPA3</i>  | GCAATGCAAGTTGCCACTCA   | ATCCTGTACCTCCGCTCACT     |
| <i>b-KLF5</i>   | GCGGGTGCTGACCATGAG     | GGTCTACGACTGAGGCAGTG     |
| <i>b-TBX3</i>   | CTGCTACTGGGAACAGTGG    | GGAGTCAGTATAGTAAATCCGTGC |
| <i>b-BCL-2</i>  | ATGTGTGTGGAGAGCGTCAA   | GGGCCATACAGCTCCACAAA     |
| <i>b-BAX</i>    | GCCCTTGCTTCAGGGTTT     | GTCCAATGTCCAGGCCATGA     |
| <i>b-BIM</i>    | TACCAGATCCCCGCTTTCA    | CGTAGCTTGGCAATCCA        |
| <i>b-CASP9</i>  | GATCAGGCCAGGCAGCTAAT   | CGGCTTGATGGTCATCCT       |
| <i>b-CCND1</i>  | GCACCTCCTCTCCAAGATGC   | GGTTGGAAATGAACCTCACG     |
| <i>b-CCNE1</i>  | GTCCCCGTGACCATTGTGTCC  | ACACCACTGATACCCTGAAACC   |
| <i>b-CCNE2</i>  | CGGGTCTGGCGAGGTTT      | AGGCGGCCAACAAATTCTTA     |
| <i>b-E2F2</i>   | GGAGGACAAGGCCAACAAAGA  | TGCCTCTGGTTGTTGAGAGG     |
| <i>b-E2F3</i>   | CCTGGAGCAGTACCTGGTGA   | CCGCTTCTCCTAGCTCCAG      |
| <i>b-E2F4</i>   | GAAGACCTGCTCCAGAACCC   | GCCTCCACTCACTGGAATCTC    |
| <i>b-E2F7</i>   | AAGTGCACGTACGGAAAGAT   | GGAGTAGACTCCTGTGGAGC     |
| <i>b-CDKN1A</i> | GAGACCGTGGTTGGGAGAC    | AAAGTCGAAGTTCCACCGCT     |
| <i>b-CDKN1B</i> | AGGACACGCATTGGTCGAT    | TTCTGAGGCCAGGCTTCTTG     |
| <i>b-RBI</i>    | CTGTGTGCACGCTTCTGTC    | GTCTCCTGAACAGCATGAGGAA   |
| <i>b-RBL1</i>   | CAAGAATGGGTGGGCCTTA    | AACTCTAGAAGGGCTGCCG      |
| <i>b-GAPDH</i>  | GGGTCATCATCTCTGCACCT   | GGTCATAAGTCCCTCCACGA     |

**Table S6.** The primers used in ChIP-qPCR analyses.

| Gene name          | Forward primer        | Reverse primer        |
|--------------------|-----------------------|-----------------------|
| 1# '-2000 to -1700 | CAGTCACTAAGTCGTGT     | GCAAATAGGTGAGAATG     |
| 2# '-1687 to -1458 | GCTTTACTCTCTCTGGCTCC  | TCAAAATATCGCTTATTGGT  |
| 3# '-1458 to -1174 | GTGCTCGTTAGAGCAT      | TTTGTTCCTTGGCTCTC     |
| 4# '-1174 to -846  | TTTCACAAAGCATCTCC     | AGGTCGCTCGTGTATTTCG   |
| 5#'-819 to -610    | TGAGGATGTAATGAAAAAGA  | GGAAGAAGATGAAAGAGAAG  |
| 6#'-576 to -367    | GAGAGGC GGAAAAGAGGTGG | TAGGGGAAAGATT CAGGCAA |
| 7#'-356 to -112    | GCGCGTACACACACACA     | GCTGAAAAGCTGCCAC      |
| 8#'-103 to +98     | GGCGGC GGGAGGAGGAGAGG | CCGCAGAGCGCACAGCCTGG  |