Supporting Information

Suntsova et al. 10.1073/pnas.1318172110

SI Methods

Tissue Samples. Tissue samples were taken from human heart, lung, bladder, testicles, brain cortex (three samples), hippocampus (16 samples taken from left and right hemispheres from eight different donors), and brain sections from one 38-y-old adult. The brain sections were taken from the left and right frontal lobe, parietal lobe, temporal lobe, occipital lobe, hippocampus, hypothalamus, thalamus, and medulla oblongata, as well as from the midbrain, pons, and left, central, and right cerebellum. All samples were taken within 24 h of death from adult (19- to 63-y-old) donors killed in road accidents.

Primary Neuronal Cell Culture. Brains of newborn Wistar rat pups were placed in ice-cold isolation solution (130 mM NaCl, 5.4 mM KCl, 20 mM Hepes, 0.4 mM KH_2PO_4, 15 mM glucose, 0.5 mM $MgSO_4$, 3 mg/mL BSA, pH 7.4). Hippocampi were removed and placed directly into fresh isolation solution. Then hippocampi were cut manually into 0.5-mm^3 pieces and were trypsinized in prewarmed isolation solution containing 4 mg/mL trypsin at 37 °C for 15 min. After trypsinization, hippocampi were centrifuged for 5 min at $900 \times G$, and pellets were resuspended in 5 mL of culture medium [modified Eagle's medium, 2 mM L-glutamine, 10% Nu-Serum (BD Biosciences), 2.5 mM glucose, 1 mM sodium pyruvate, 15 mM Hepes]. To remove tissue debris, the cell suspension was kept on the bench for 2 min, and 1 mL of the upper fraction was transferred to a fresh tube for future preparation. Then 1 mL of fresh culture medium was added to the initial cell suspension. The procedure was repeated four times; all the aliquots were combined and centrifuged for 5 min at 1,800 rpm. The pellet was resuspended in 1 mL of fresh culture medium, and the concentration of cells was determined using a hemocytometer. Cells were plated on poly-D-lysine–coated coverslips in 24-well plates at a density of ∼250,000 cells per well. Culture medium was replaced after 1 h.

Genomic DNA and RNA Extraction and cDNA Synthesis. Genomic DNA was isolated using the Wizard Genomic DNA Purification Kit (Promega), RNA was extracted using the SV Total RNA Isolation System (Promega), and first-strand cDNA synthesis was conducted using the Mint reverse transcription kit (Evrogen), according to the manufacturers' recommendations.

PCR Amplification. For the amplification of upstream gene regions and the bisulfite conversion of DNA, we used the Encyclo PCR Kit (Evrogen). For real-time, quantitative PCR (qRT-PCR), we used the qPCRmix-HS SYBR mix (Evrogen). Real-time RT-PCR was performed on a MxPro3000 thermocycler (Stratagene). Primer sequences are listed in Table S1.

DNA Cloning. We used genomic DNA isolated from human placenta (∼40 ng per reaction) for upstream PCR reactions. The PCR products were ligated into pGEM-T Easy vectors (Promega) and subcloned into a pGL3-basic vector (Promega) and a pDsRed-Express-N1 vector (Clontech). For molecular cloning, we used the XL1-Blue *Escherichia coli* strain type recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′proAB lacIqZΔM15 Tn10 (Tetr)], which is deficient in both site-specific recombination and DNA adenine methyltransferase methylation.

Lentiviral Vector Construction. A 42-bp linker containing Bsp120I and MluI restriction sites was cloned by AvaI and HpaI into the lentiviral vector p156RRLsinPPTCMVGFPPRE (a gift from Alon Chen, the Weizmann Institute of Science, Rehovot, Israel). An expression cassette containing DsRed fluorescent protein under the control of the *PRODH* upstream regulatory sequence was subcloned into the resulting vector.

Cell Transfections and Luciferase Assay. Transfections were performed in 24-well plates using the Unifectin-56 transfection reagent according to the manufacturer's recommendations. For each well, 0.5 μg of DNA was used. pGL3-based reporter constructs carrying tested gene upstream regions and the firefly luciferase gene were mixed in a ratio of 10:1 with a pRL-TK vector used as an internal control for transfection. The values obtained for firefly luciferase were normalized to the values for the Renilla luciferase. The pRL-TK vector has the *Renilla reniformis* luciferase gene under the control of the upstream HSV-thymidine kinase promoter. For overexpression of *SOX2* and *NF-κB*, we used plasmid vectors pMXs-hSOX2 (Addgene; kindly provided by Maria Lagarkova, the Vavilov Institute of General Genetics, Moscow) and phNFKB1 (kindly provided by Alexander Belyavsky, the Engelhardt Institute of Molecular Biology, Moscow), respectively. In these constructs, the genes are placed under the control of the standard CMV promoter. Luciferase activity was measured 48 h after transfection using the Dual Luciferase Reporter Assay (Promega) and a GENios Pro luminometer (Tecan). All experiments and measurements were done at least in quadruplicate.

Microscopy. On the third day after transfection, cells were fixed and mounted on the microscopic slides using ProLong Gold Antifade reagent (Life Technologies). Images were taken with an LSM 5 Life confocal microscope (Zeiss) and were pseudocolored using ImageJ image-analysis software (1). The cell types (neurons, astrocytes, and microglia) were distinguished morphologically by shape.

Microarray Hybridization. Microarray hybridization was performed with the HumanHT-12 v4 Expression BeadChip Kit (Illumina) according to the manufacturer's recommendations. Microarray hybridization was carried out at the Genoanalytica core facilities.

In Vitro DNA Methylation. pGL3-based reporter constructs were methylated in vitro using M.Sss I CpG-methyltransferase (SibEnzyme) according to the manufacturer's protocol. After incubation for 3, 10, or 30 min, reactions were heat inactivated, and DNA was purified for use in transfection experiments.

Bisulfite Sequencing. Bisulfite conversion of genomic DNA was performed using the EpiTect Bisulfite Kit (Qiagen) followed by amplification with the EpiTect Whole Bisulfitome Kit (Qiagen) to increase the quality and stability of the converted DNA. DNA then was amplified by nested PCR using the PCR primers listed in Table S1. PCR products were cloned into pGEM-T Easy vectors and sequenced directly. Sequence data were analyzed using Biq Analyzer software.

Methylation-Sensitive, High-Resolution Melting Assay. Bisulfite-converted and subsequently amplified DNA corresponding to the $hsERV_{PRODH}$ LTR was purified and used as a template for the methylation-sensitive, high-resolution melting (MS-HRM) assay. DNA was amplified with specific primers for MS-HRM using the qPCRmix-HS SYBR. PCR products were melted following the standard melting protocol for the MxPro3000 thermocycler (Stratagene).

1. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.

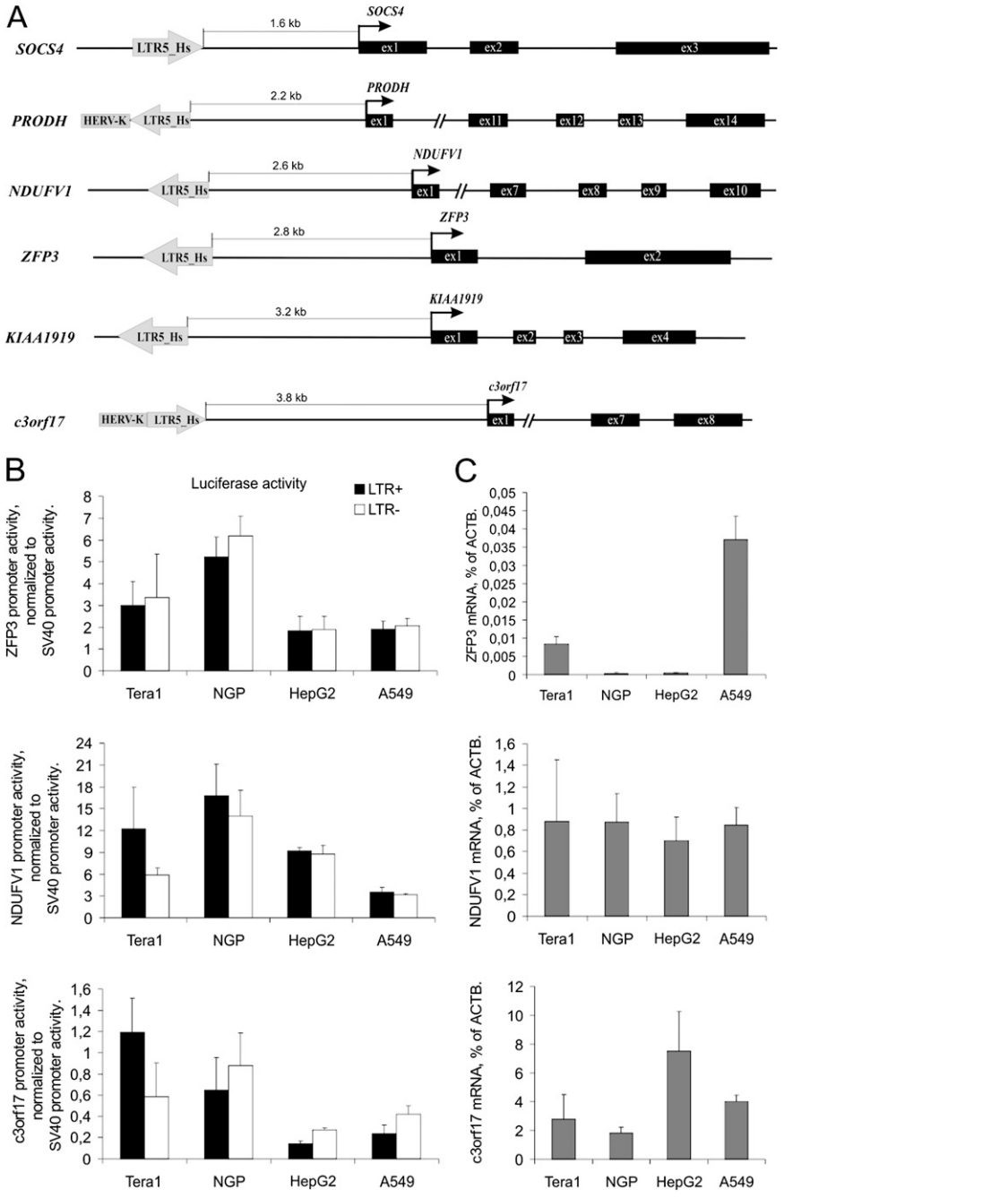

Fig. S1. Comparison of luciferase reporter assay data with transcriptional activities of the endogenous (genomic) gene copies. (*A*) Schematic representation of hsERV elements located close to transcription start sites of known genes. Gray arrow, HsERV LTR sequence; gray bar, hsERV internal region; black arrow. gene transcription start site; black bar, exon. (*B*) Dual luciferase reporter gene assay used to assess transcriptional activities of LTR^{+} and LTR^{-} constructs in Tera-1, NGP127, HepG2, and A549 cell lines. LTR^{+} and LTR^{-} transcriptional activities were normalized to the SV40 promoter activity in every experiment. (*C*) mRNA levels of *ZFP3*, *NDUFV1*, and *c3orf17* genes in Tera-1, NGP127, HepG2, and A549 cell lines. β-Actin (*ACTB*) was used for qRT-PCR normalization.

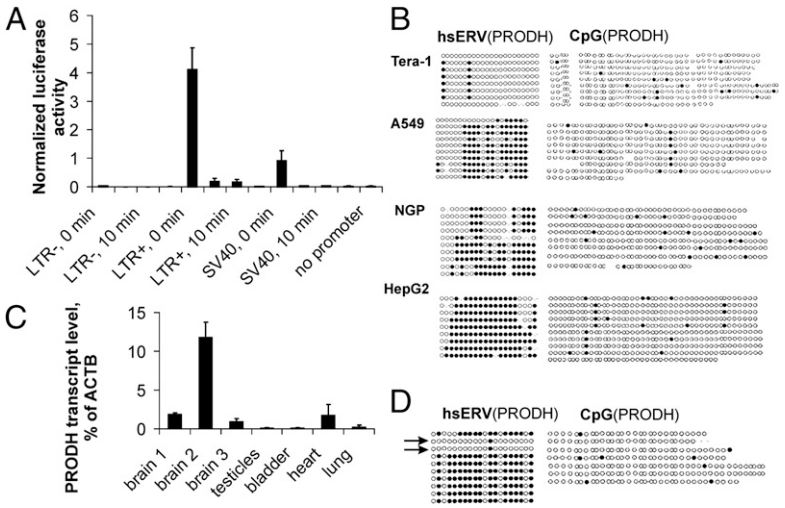

Fig. S2. DNA methylation and *PRODH* promoter activity. (*A*) Transfection of Tera-1 cells with in vitro-methylated reporter constructs. The LTR^{+} and LTR^{-} and the pGL3-SV40 reporter construct were methylated in vitro for 0, 3, 10, and 30 min. (*B* and *D*) Methylation patterns of hsERV$_{PRODH}$ (*Left*) and CpG$_{PRODH}$ (*Right*) in Tera-1 and A549 cells (*B*) and in brain sample number 1 (*D*) were determined by bisulfite sequencing. Arrows indicate hypomethylated alleles of hsERV$_{PRODH}$. Black circle, methylated CG dinucleotide; open circle, unmethylated CG dinucleotide. (*C*) Relative *PRODH* transcript levels in the human tissues (normalized to *ACTB*). Data show means ± SD of three independent experiments.

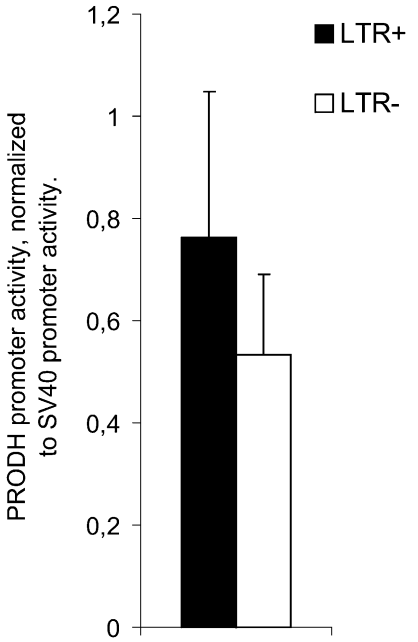

Fig. S3. Normalized *PRODH* promoter activity in NT2/D1 cells. The NT2/D1 cell culture was transfected with the LTR^{+} and LTR^{-} luciferase reporter constructs. *PRODH* upstream region promoter activities were normalized to the SV40 promoter activity in the same experiments.

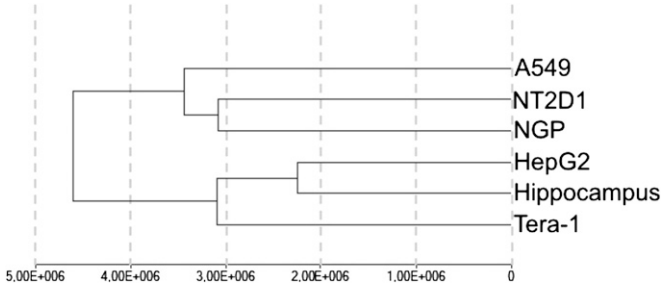

Fig. S4. Cluster analysis of microarray gene-expression data in cell cultures of Tera-1, A549, HepG2, NGP127, and NT2D1 cells and in a human hippocampus tissue sample. The scale defines the relative differences in gene-expression profiles among the cell cultures.

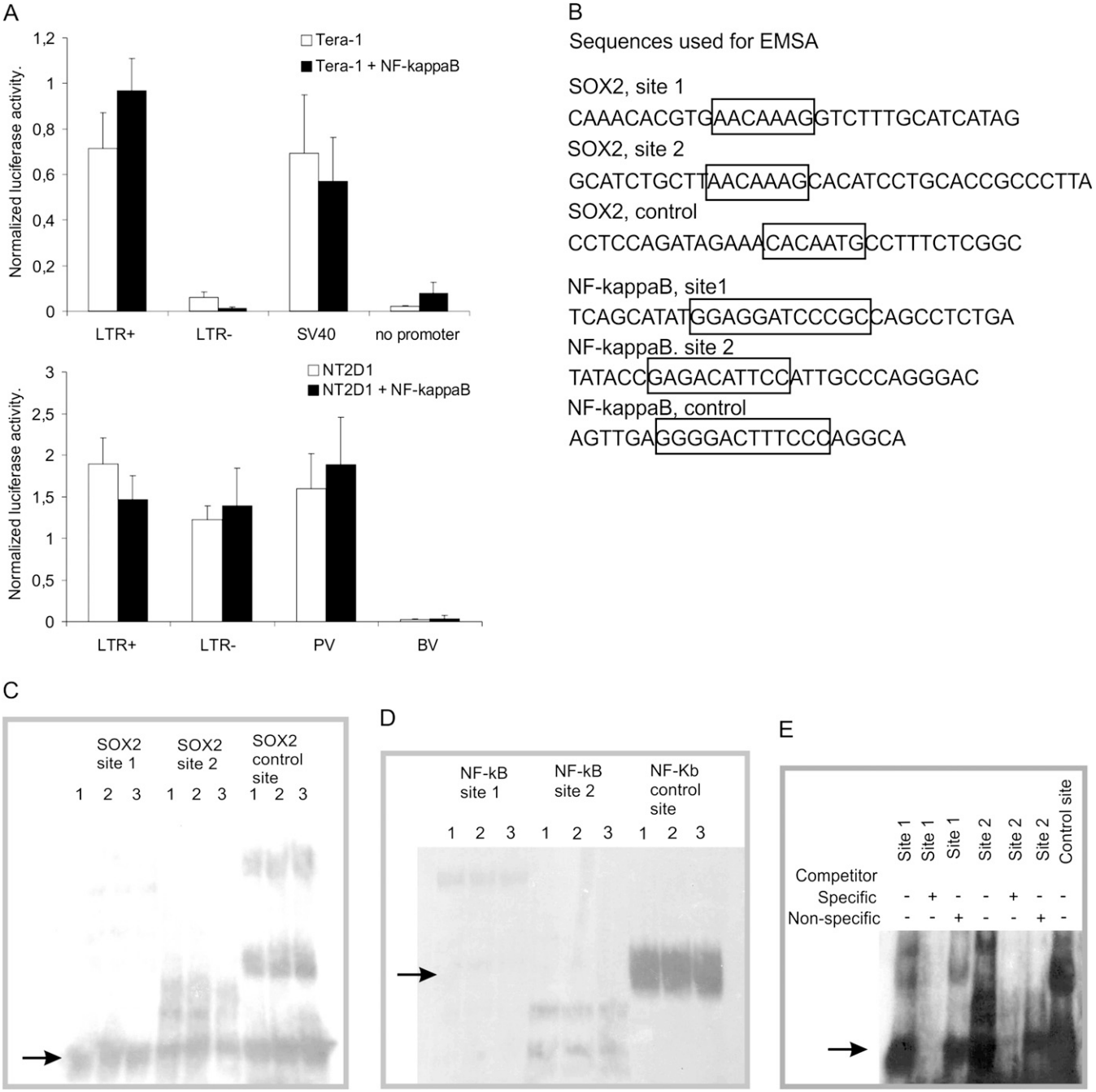

Fig. S5. Regulation of hsERV$_{PRODH}$ enhancer activity by transcription factors. (*A*) Effect of *NF-κB* gene overexpression on hsERV$_{PRODH}$ enhancer activity. Tera-1 (*Upper*) and NT2/D1 (*Lower*) cells were transfected with luciferase reporter constructs alone or in combination with a construct carrying the human *NF-κB* gene under the control of the CMV promoter. *Firefly* luciferase activity was normalized to *Renilla* luciferase activity. (*B*) Structures of oligonucleotides used in EMSA tests. The double-stranded oligonucleotides used here had been experimentally verified previously (controls) or were predicted to be SOX2 and NF-κB recognition sites. SOX2 sites 1 and 2 and NF-κB sites 1 and 2 were predicted by software to be within the hsERV$_{PRODH}$ sequence. For positive controls we took oligonucleotides containing canonical SOX2 (1) and NFKB1 (NFκB gel-shift oligonucleotides; Santa Cruz Biotechnology) transcription factor-binding sites. (*C*) EMSA for SOX2 putative binding sites 1 and 2 and for the control SOX2-binding site, with nuclear extract of Tera-1 cells, each with three replicates. (*D*) EMSA for NF-κB putative binding sites 1 and 2 and for the control NF-κB–binding site, with nuclear extract of Tera-1 cells, each in three replicates. (*E*) Competition assay. EMSA for the SOX2 putative binding sites 1 and 2 and for the control SOX2-binding site, with nuclear extract of Tera-1 cells with or without a 20-fold excess of nonradiolabeled competitor double-stranded oligonucleotides. As a specific competitor, we used control SOX2-binding oligonucleotide, and as a nonspecific competitor we used control NF-κB–binding oligonucleotide.

1. Card DA, et al. (2008) Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol Cell Biol* 28(20):6426–6438.

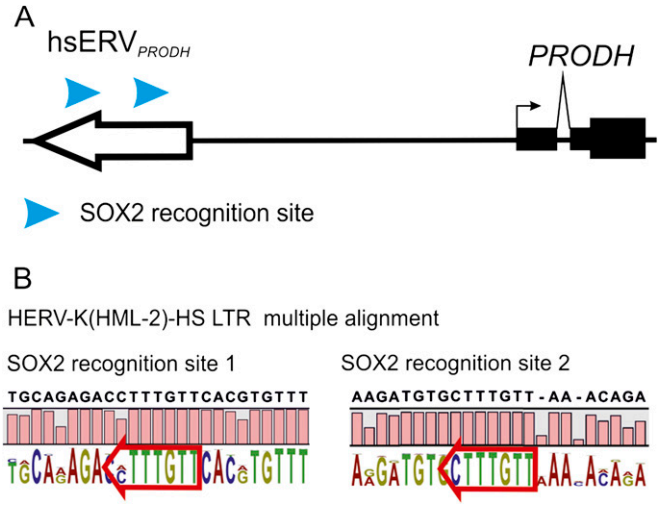

Fig. S6. (*A*) Schematic representation of SOX2-binding sites in the *PRODH* upstream regions. (*B*) Multiple sequence alignment of SOX2-binding sites within the HERV-K(HML-2) family LTR sequences. SOX2 core recognition sites (arrows) are in the opposite orientation relative to the LTR direction.

Table S1. Sequences of oligonucleotides used in this study

Forward oligonucleotide	Reverse oligonucleotide
Primers for amplification of promoter regions	
KIAA1919-LTR TTACGCGTGGAGGATTGATTGA	KIAA1919 TTCTCGAGTCAAGCCGGAAGTT
KIAA1919 TTACGCGTCATCACCAATGTCA,	KIAA1919 TTCTCGAGTCAAGCCGGAAGTT
ZFP3-LTR TTACGCGTTTCCAATGTGCAGT,	ZFP3 TTAGATCTGGCCACTTCCGGTA,
ZFP3 TTACGCGTAATCACGATTTAGG,	ZFP3 TTAGATCTGGCCACTTCCGGTA,
SOCS4-LTR AGGTACCACCCAAACTTGTGGGGAAAA,	SOCS4 ATAGATCTGAGCGAGCCGAAGC,
SOCS4 ATACGCGTATCTTCTAACAGTC,	SOCS4 ATAGATCTGAGCGAGCCGAAGC,
PRODH-LTR ATACGCGTTGTAGACGTACTCG,	PRODH ATAGATCTCTCTCCGTGAAGAC,
PRODH ATACGCGTTGGGTAACACAGCC,	PRODH ATAGATCTCTCTCCGTGAAGAC,
PRODH-UTR TAGATCTTTTGAGAAGCCAGTGGGCCTTG,	PRODH-UTR TAAGCTTCAGAGCtATGGCGGGACGGCGGTA,
PRODH-Chimp ATAGATCTTTCACAAGCACCCTGGCA	PRODH-Chimp ATAGATCTTTCACAAGCACCCTGGCA
NDUFV1-LTR ATACGCGTTGCTGCAGGCTTGG,	NDUFV1 ATAGATCTGGCTGAGGAACTGG,
NDUFV1 ATACGCGTGGGACGGTTGTAGG,	NDUFV1 ATAGATCTGGCTGAGGAACTGG,
C3orf17 –LTR GGTACCGCCATGATGACGACGGTG,	C3orf17 GGTACCCGTCCAGCGACCGGAGG;
C3orf17 GGTACCAGGCTGGTGGAATCCTGACT,	C3orf17 GGTACCCGTCCAGCGACCGGAGG;
Primers for SOX2 amplification	
SOX2T7-for ACACACATAATACGACTCACTATAGGGTGGTACGGGAAATCACAAG	
T50 $(T)_{50}$AACTTGTTTATTGCAGCTTATAATGG	
Primers for real-time PCR	
PRODHq GCAGAGATCGGCTATGAGGAC	PRODHq GGAAGCTGATCTGGTCACACA
NDUFV1q GACATCGTGAAAGCCATCG	NDUFV1q CTGATGCTGCTGGGCAAAC
KIAA1919q GTGGCTCTGTGATTGGTGGA	KIAA1919q CGATAACAGCATAAGCCCACA
C3orf17q GGCAGTTGTATGGTGCAGGA	C3orf17q GGCAGTTGTATGGTGCAGGA
SOCS4q CTCCAGTGACTTCCGTTTGTG	SOCS4q TCCACTCCACACATAACCGTC
ZFP3q TGAGTTTGGAGCAGGATGTGA	ZFP3q TGCACTAACTCCCTCTGTTGTATC
β-Actin GAGCGGGAAATCGTGCGTGACATT	β-Actin GATGGAGTTGAAGGTAGTTTCGTG
Primers for bisulfite sequencing	
bisPRODH-out GGTTTAGTTTGGGTTTAGAGGTTTG	bisPRODH-out TACCTTCACCCTAAAAAAAAACCTCCA
bisPRODH-in GTGGAGATAGGGTTTGGTTGTGTTATTTAGTG	bisPRODH-in GTGGAGATAGGGTTTGGTTGTGTTATTTAGTG
bisPRODH-CPG-out GTGATAATTTTGTATTTTTATAAGTGGTA	bisPRODH-CPG-out AAAATACCAAAATACTTATAAAATCTAAAC
bisPRODH-CPG-in TTGGTTTTTTTTGTTTTAGGAGATGTTTG	bisPRODH-CPG-in TATCAAAAATATAAAATAAAACCTCTAAAAC
bisSOCS4-out ATATTTATGTTTATTTTAGTTTAAGTTTT	bisSOCS4-out CAATAAATATCCTATCCAAAACCTTCTAAAC
bisSOCS4-in ATTTGTGGGGAAAAGTAAGAGAGA	bisSOCS4-in AATACAAAAAAACTATTAAAAAATACC
bisKIAA1919-out AAGTTATGATTATGTTATTGTATTTTAGT	bisKIAA1919-out AATATAAACTTAAAATTTTCTTCTTTCT
bisKIAA1919-in TTGGGTGATTGATTAAGATTGTT	bisKIAA1919-out TTCTTCCTTTCTTCTTTCTCTATA
Primers for MS-HRM	
Ms-hrm 1 TTTTATGTGATAGTTTGAAATATGG	Ms-hrm 1 TACTAATTCTCCTCAACACAAACC
Ms-hrm 2 TGTTTTTGGGTAATGGAATG	Ms-hrm 2 ATCCCACCTCCAACCCTAA
Oligonucleotides for EMSA	
EMSASox2-1 CAAACACGTGAACAAAGGTCTTTGC	EMSASox2-1 CTATGATGCAAAGACCTTTGTTC
EMSASox2-2 GCATCTGCTTAACAAAGCACATCCTGC	EMSASox2-2 TAAGGGCGGTGCAGGATGTGCTTTGTTAAGC
EMSASox2-1mut CAAACACGTGAAATTACGTCTTTGC	EMSASox2-1mut CTATGATGCAAAGACGTAATTTC
EMSASox2-2mut GCATCTGCTTAAATTACCACATCCTGC	EMSASox2-2mut TAAGGGCGGTGCAGGATGTGGTAATTTAAGC
EMSASox2 control CCTCCAGATAGAA	EMSASox2 control GCCGAGAAAGGCATTGTGTTTCTATCTGGAGG
EMSANfkb-1 TCAGCATATGGAGGATCCCGCC	EMSANfkb-1 TCAGAGGCTGGCGGGATCCTCCAT
EMSANFKB-2 TATACCGAGACATTCCATTGCCC	EMSANFKB-2 GTCCCTGGGCAATGGAATGTCTC
EMSANFKB control AGTTGAGGGGACTTTCCC	EMSANFKB controlTGCCTGGGAAAGTCCCCTC

Other Supporting Information Files

Dataset S1 (XLS)
Dataset S2 (XLS)
Dataset S3 (XLS)