Supplementary Information for

Regulation of stem cell function and neuronal differentiation by HERV-K via mTOR pathway

Tongguang Wang, Marie Medynets, Kory R. Johnson, Tara T. Doucet-O'Hare, Brianna DiSanza, Wenxue Li, Yadi Xu, Anna Bagnell, Richa Tyagi, Kevon Sampson, Nasir Malik, Joseph Steiner, Alina Hadegan, Jeffrey Kowalak, James O'Malley, Dragan Maric, Avindra Nath[#]

[#]correspondence:

Avindra Nath,
Email: natha@ninds.nih.gov

SI Methods:

RNA-Seq analysis

RNA-Seq analysis was performed on the iPSC lines NCRM1, NCRM5, and ND2.0. RNA-Seq analysis was also conducted on NSCs that were derived from NCRM1 and BC1 iPSCs and the H14 embryonic stem cell line (1) and on neurons differentiated from these NSCs at the following time points: days 3, 5, 7, and 10. RNA-Seq data from the XCL1 NSC line (XCell Science, Novato, CA) was also used as part of the RNA-Seq data analysis. NSCs were derived from the pluripotent stem cell lines following the protocol described by Swistowski et al.(2) and neurons differentiated according to Efthymiou et al. (3). Cells were grown in one well of a 6-well dish, samples processed at indicated time points, and total RNA extracted with the RNeasy Mini Kit (Qiagen, Redwood City, CA). RNA concentration and integrity were assessed with a BioAnalyzer (Agilent, Santa Clara, CA). RNA sequencing library was generated from the RNA by using the TruSeq SR RNA sample prep kit (Illumina, San Diego, CA) and single read sequenced on HiSeq 2000 (Illumina) for 50 cycles. Post CASAVA deplexing, fastq files

generated per library were import into the CLCbio Genomics Workbench (https://digitalinsights.qiagen.com) and adaptor clipped, QC inspected, quality trimmed, filtered, and mapped. Per quality trimming, 15bp from the 5' end of each read was globally removed along with 1bp from the 3' end. Bases with a call accuracy < 95% were also trim-removed. Per filtering, reads having more than two ambiguities were discarded along with those having read length <15bp post quality trimming. Per mapping, the "RNA-Seq" tool was used to locally align reads by library against the human genome (GRCh38/hg38) without masking; keeping only those mapped reads with >= 80% reference identity for >=80% read length. Post mapping, the number of reads falling in a RepeatMasker annotated "HERV" region (Library 20120124) were enumerated and the counts converted into "Reads Per Kilo base of transcript per Million mapped reads" ("RPKM") expression values. RPKM expression values were next organized by HERV region across libraries and imported into R (https://cran.r-project.org/). In R, RPKM expression values were pedestalled by 2, Log2 transformed, filtered to remove regions not having at least one transformed value >1, then quantile normalized. Normalized values for remaining regions were then used to assess for and remove library-level outliers via Tukey box plot, covariance-based PCA scatter plot, and Pearson correlation heat map. For libraries not removed as outliers, normalization of RPKM expression values was repeated and the new normalized expression noise modeled using LOWESS (CV~mean). Based on the LOWESS fit, regions not having at least one normalized value >=2.25 were considered noise-biased and subsequently discarded as such. For regions not discarded, expression was floored to 2.25 if < 2.25 then subject to one-factor ANOVA testing under BH FDR MCC (BH: Benjamini-

Hochberg, FDR: False discovery rate, MCC: Matthews Correlation Coefficients) condition using sample type as the factor. Regions having a corrected $P < 0.05$ by this test were subset and deemed to be those having differential expression across the sample types tested. For those regions subset only, the TukeyHSD post-hoc test was applied to identify which regions between sample types are differentially expressed ($P < 0.05$ and an absolute difference of means >=2).

Global DNA methylation analysis

Induced neural stem cells and their corresponding iPSCs were lysed and genomic DNA was purified using a DNA purification kit (Clontech, Mountain View, CA) following the manufacturer's instructions. LINE-1 methylation level was used to represent the global DNA methylation level and was detected using the Global DNA methylation LINE-1 kit (Active Motif, Carlsbad, CA) per the manufacturer's instructions.

Liquid chromatography-mass spectrometry

To determine the possible binding proteins with HML-2 Env, HML-2 Env magnetic beads were made by conjugating protein G magnet beads (50 μl) with anti-Flag mAb (5 μg, Sigma) for 30 mins. The beads were then incubated with cell lysates of HEK293 cells transfected with HML-2 Env FlagTag plasmid for 1 hour at room temperature, followed by incubation with cell lysates of iPSC. Immunoprecipitated samples were processed using an on-bead digestion (OBD) protocol (4) with minor changes. The OBD protocol yields 4 fractions per sample based on increasingly vigorous conditions. Briefly, proteins immobilized on magnetic beads were resuspended in 200 μL 0.1M NH_4HCO_3, pH 8.0. Disulfide bonds were reduced by addition 1.0 μL 1M dithiothreitol (DTT, final concentration = 5 mM) and incubating at 50^oC for 30 min. The supernatant was

decanted from the beads and transferred to a 1.5 mL microfuge tube (fraction 1) and a fresh aliquot of 200 µL 0.1M NH_4HCO_3, pH 8.0 was added to the magnetic beads. Thiol moieties were alkylated by addition of 9.0 µL 0.29M iodoacetamide (IAM, final concentration = 12.5 mM) to both fraction 1 and the resuspended magnetic beads and incubating at room temp for 30 min in the dark. Proteins were digested by adding 10 µL trypsin (10 ng/µL) and incubating at 50^{o}C for 4 hr. The supernatant was decanted from the magnetic beads and transferred to a 1.5 mL microfuge tube (fraction 2). The magnetic beads resuspended in 200 µL 2M urea, 0.1M NH_4HCO_3, pH 8.0. Proteins were further digested by adding 10 µL trypsin (10 ng/µL) to fractions 1, 2, and the magnetic beads, and incubating at 37^{o}C overnight. The supernatant was decanted from the magnetic beads and transferred to a 1.5 mL microfuge tube (fraction 3). The magnetic beads were resuspended in 200 µL 80% CH_3CN, 0.1% trifluoroacetic acid (TFA) and sonicated for 30 min. The supernatant was decanted from the magnetic beads and transferred to a 1.5 mL microfuge tube (fraction 4) and dried in a vacuum centrifuge.

Fractions 1 – 3 were acidified by the addition of 1.0 µL TFA, while fraction 4 was resolubilized in 200 µL 5% CH_3CN, 0.1% TFA. All fractions were desalted by solid phase extraction using Targa C18 microspin columns (PN: SEM SS18R; The Nest Group, Inc., Southborough, MA) according to the manufacturer's instructions.

Nanoflow LC-MS/MS was performed using an Easy nLC1000 UHPLC system directly coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher). An integrated autosampler was used to load samples at 5 µL/min onto a nanoViper trap column (75 µm x 20 mm, Acclaim PepMap 100 C18 resin, 3 µm particle size, 100Å pore size).

Reversed phase HPLC was performed using an EasySpray C18 column (75 µm ID x 250 mm; Acclaim PepMap 100 C18 resin, 2 µm particle size, 100Å pore size). Mobile phases were Buffer A: H_2O:DMSO:HCO_2H (95:5:0.1; v/v/v) and Buffer B: CH_3CN:DMSO:HCO_2H (95:5:0.1; v/v/v). After 5 min of isocratic flow (200 nL/min, 2%B) a linear gradient from 2% to 35 %B was developed over 66 min, 35-80% B over 4 min, and isocratic flow at 80% for 5 min.

LC-MS/MS experiments were performed using data dependent acquisition with dynamic exclusion enabled (exclusion width = ± 10 ppm, repeat count = 2, repeat duration = 15 sec, exclusion duration 22.5 sec). MS1 scans were acquired in the Orbitrap mass analyzer (resolution 15,000 at m/z 400). The five most intense peptide ions from each survey scan were sequentially isolated and fragmented in the linear ion trap (isolation width = 2Da, normalized collision energy = 35%, activation Q = 0.25, AGC target = $1.0 \times 10_4$).

Charge state exclusion enabled (1+, ≥ 5+).

Raw data files were analyzed using PEAKS Studio 8.5 bioinformatics package (Bioinformatic Solutions, Inc., Waterloo, ON). MS/MS spectra were searched against a custom database constructed from the SwissProt portion of the UniProt database (release 2019_06). The database included all HUMAN sequences in SwissProt (20,471 sequences) and was augmented with reagent protein sequences (ASPN_PSEFR, CLOS_CLOHI, CTRA_BOVIN, CTRB_BOVIN, CTRC_BOVIN, LYSC_PSEAE, PNGF_ELIMR, and TRYP_PIG, SPA_STAA8, SPA_STAAU, SPG1_STRSG, SPG2_STRSG, GST26_SCHJA, 3Xflag peptide).

Relevant search database parameters were: enzyme = trypsin/P, maximum missed cleavages = 2, fixed modifications = carbamidomethyl-Cys (C), variable modifications = N-acetyl (Protein), pyro-Glu (Q), and oxidation (M), precursor ion mass tolerance = ±10 ppm, product ion mass tolerance = 0.6 Da, mass values = monoisotopic. Only high confidence peptide sequence assignments were accepted. Peptide validation False Discovery Rate < 1%. Protein ID based on ≥ 2 sibling peptides.

1. J. Beers *et al.*, Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nat Protoc* **7**, 2029-2040 (2012).
2. A. Swistowski *et al.*, Efficient generation of functional dopaminergic neurons from human induced pluripotent stem cells under defined conditions. *Stem Cells* **28**, 1893-1904 (2010).
3. A. Efthymiou *et al.*, Functional screening assays with neurons generated from pluripotent stem cell-derived neural stem cells. *J Biomol Screen* **19**, 32-43 (2014).
4. M. J. Berberich *et al.*, "Development of an On-Bead Digestion Procedure for Immunoprecipitated Proteins" in Sample Preparation in Biological Mass Spectrometry, A. R. Ivanov, A. V. Lazarev, Eds. (Springer Netherlands, Dordrecht, 2011), 10.1007/978-94-007-0828-0_7, pp. 109-124.

Supplementary Data

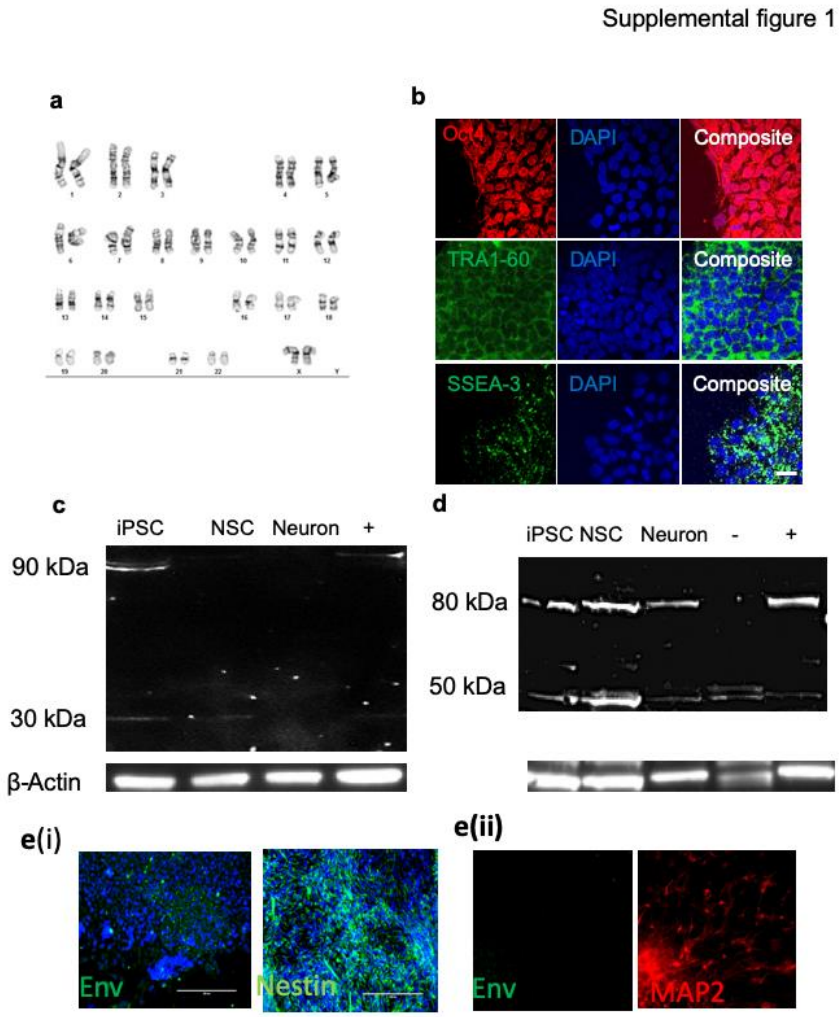


Figure S1. **Characterization of iPSC lines and detection of HERV-K protein expression by Western blot analysis.** iPSC lines were generated by transfecting CD34+ cells with Sendai virus containing Yamanaka factors.

(a) Representative karyotyping of iPSC line PAU showed normal karyotype.

(b) Pluripotent stem cell markers, Oct4, TRA1-60 and SSEA-3 are shown by immunostaining. DAPI (blue) was used to stain the nuclei. Scale bar: 20 µm.

(c) HERV-K Env and

(d) Gag protein expression in iPSC, neural stem cells (NSC) and neurons during differentiation was determined by Western-blot assays. Positive control (+) used recombinant protein. Negative control represents cell lysates from a mouse fibroblast cell line (-). β- actin was used as a loading control.

(e) (i) decreased HERV-K Env in NSC (immunostained for nestin; green) and (ii) lack of HERV-K Env in differentiated neurons (immunostained for MAP-2; red). DAPI (blue) was used to stain the nuclei.

Supplemental figure 2

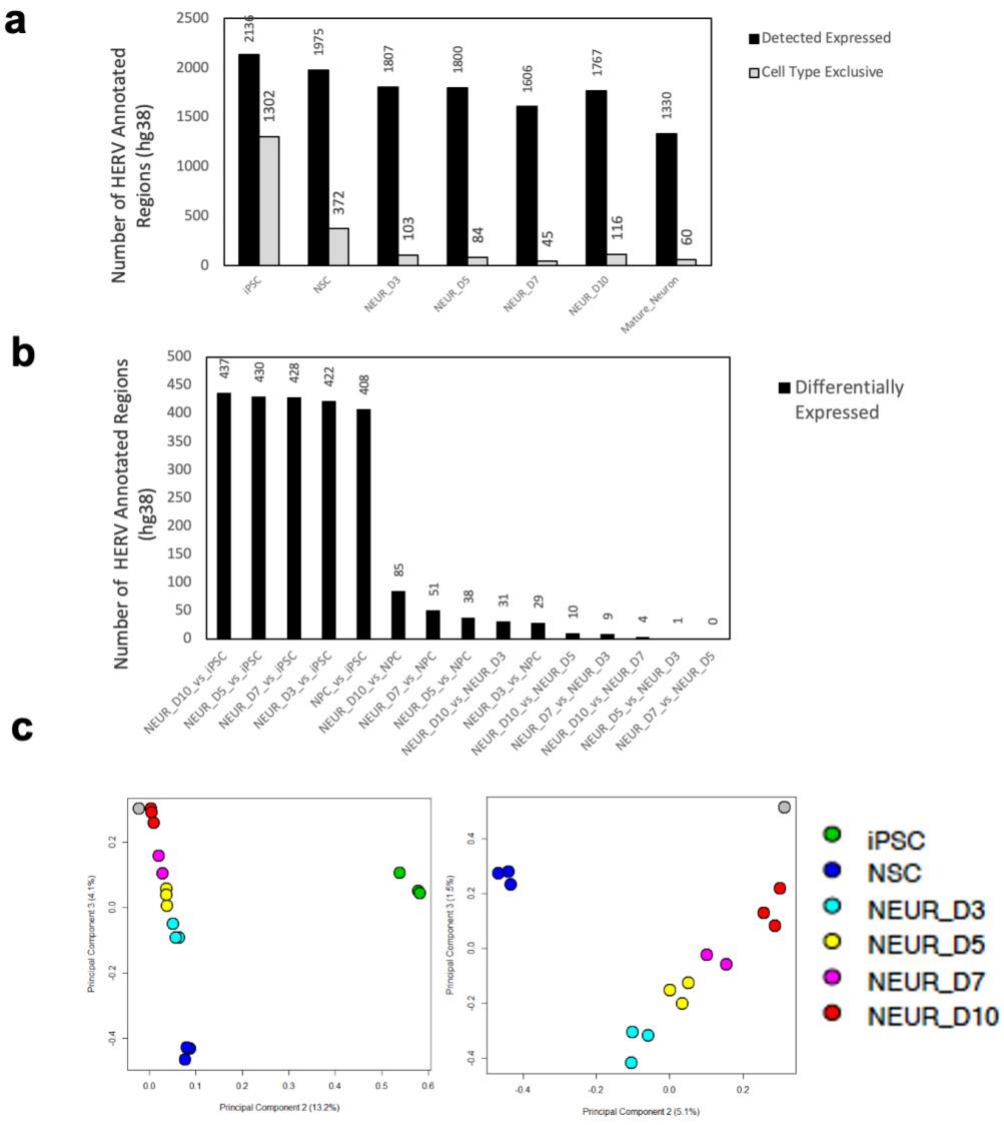


Figure S2. Differential expression of HERVs in iPSC, NSC and neurons. (a) Following RNA seq analysis, 3,698 HERV annotated regions (RepeatMasker) of the Human Genome (hg38) were observed expressed in at least one cell type profiled with maximal expression in iPSC.

(b) 541 of the HERV regions observed expressed were found to have differential expression between one or more cell types (ANOVA corrected $P < 0.05$, TukeyHSD $P < 0.05$, absolute mean expression difference >=2X.

(c) Inspection of cell type relationships by cov PCA using expression for the 541 HERV regions reveals good grouping within and between cell types; separating by differentiation stage and time.

Supplemental figure 3

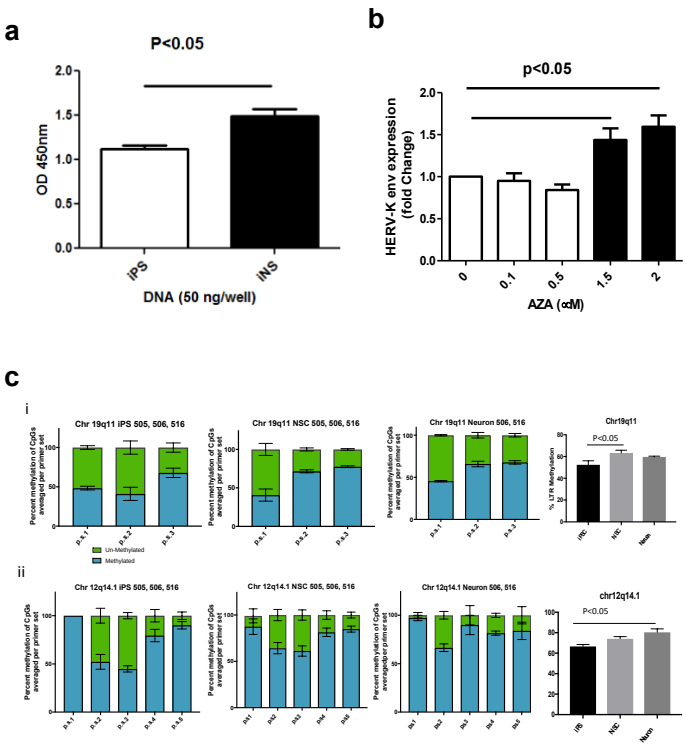


Figure S3. **HERV-K LTR methylation during neural differentiation**.

Global DNA methylation-LINE-1 kit was used to quantitate 5-methylcytosine levels in genomic DNA.

(a) global DNA methylation level was increased significantly in neural stem cells (iNS) differentiated from iPSCs (P<0.05).

(b) human neural stem cells treated with 5-aza-2′-deoxycytidinenon (AZA), which causes non-specific DNA demethylation, increased HML-2 Env expression after 24

hours. Data represent mean ± SEM from three independent experiments. Statistical analysis was performed by one-way ANOVA followed by Bonferroni's test.

(c) Methylation in HML-2 LTR at loci on chromosomes 12q14.1 and 19q11. Graphs displaying percentage methylation for each primer set (p.s.) covering an LTR at either locus (i) 12q14.1 or (ii) 19q11 of iPSCs, NSCs or neurons from three donors (505, 506 and 516). Green denotes percentage of CpGs unmethylated and blue denotes percentage of CpGs with methylation. Summarized graphs of both loci showed that the methylation level increased significantly in NSCs or neurons compared to iPSCs. Data represents mean ± SEM.

Supplemental figure 4

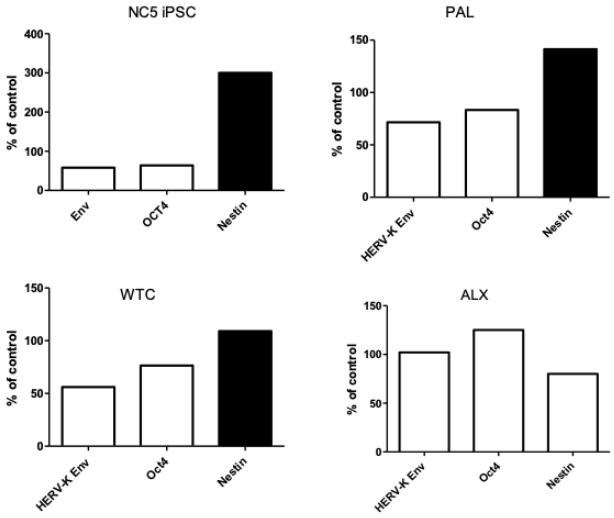


Figure S4. **<u>Effect of HML-2 env siRNA on iPSC differentiation</u>**. Inhibition of HML-2 *env* using siRNAs, on Oct3/4 and nestin transcripts in four different iPSC lines (NC5, PAL, WTC and ALX). In three of the lines where *env* was inhibited, there was increased nestin production at 24 hours.

Figure S5. **Plasmid map for HML-2 *env***. The consensus sequence of HML-2 *env* (coEnv) was inserted under a CMV promoter. The protein was linked to a FLAG tag. The consensus sequence of HML-2 *gag* was also used instead of HML-2 *env* for the plasmid construct of HML-2 *gag*.

Supplemental figure 6

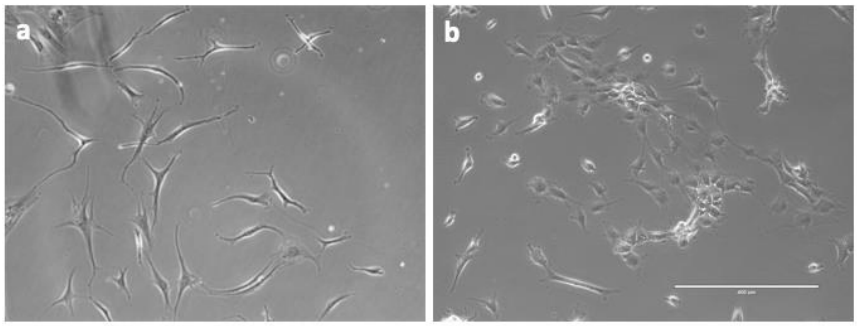


Figure S6. **Rhesus neural progenitor cells derived from fibroblasts.**

(a) Rhesus fibroblasts were transduced with Sendai virus containing Yamanaka factors and then incubated in neural induction medium for 2 weeks to produce (b) neural stem cells.

Supplemental figure 7

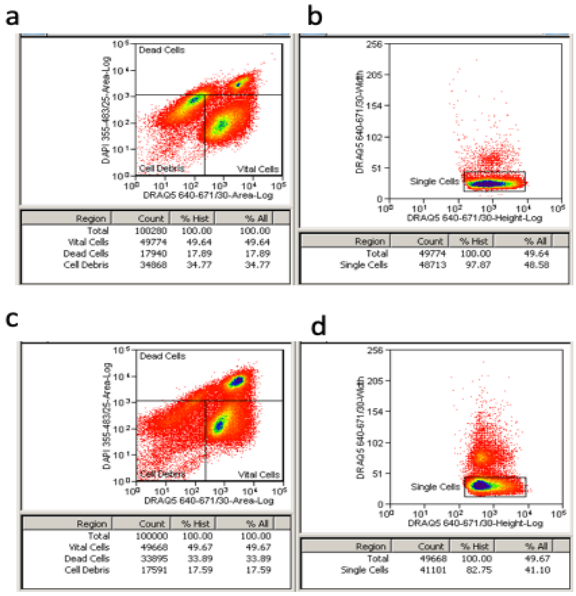


Figure S7. **FACS gating strategy.** The gates used for single cell gating of vital iPSCs for assaying HML2 Env and Tra1-60 expression are shown with DRAQ5-Area Log versus DAPI-Area Log, and DRAQ5-Height versus DRAQ5-Width bivariate plots (a and b, control staining without primary antibodies; c and d, with primary antibodies).

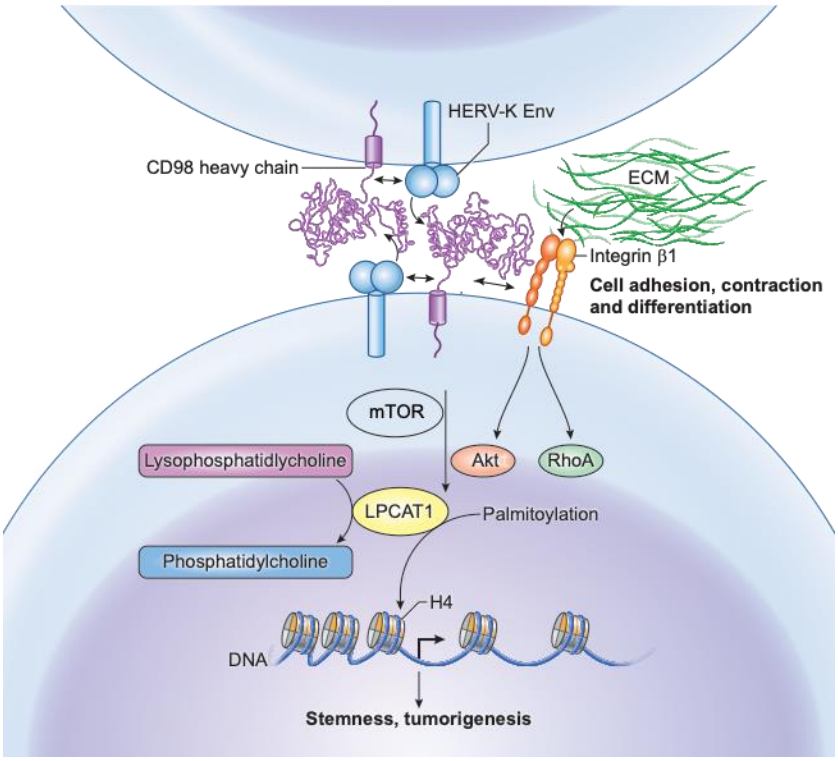


Figure S8. <u>Supplemental Schematic Figure: Mechanism underlying HERV-K subtype HML-2 Env function in human pluripotent stem cells.</u> Two pluripotent stem cells are shown in close vicinity. The HML-2 Env protein is expressed on their surfaces that binds to the CD98 heavy chain on the opposite cell. In the extracellular region, CD98 heavy chain binds to beta-1 integrin which in turn interacts with the extracellular matrix. Interactions between HML-2 Env and CD98 heavy chain also lead to activation of a cascade of secondary signaling molecules. This includes the mTOR pathway which is further linked to activation of LPCAT1. This enzyme can modulate the chromatin and maintain stemness by palmitoylation of histones. Silencing of this pathway is critical for differentiation of the pluripotent stem cells into neural progenitor cells and neuronal differentiation. Conversely, persistent activation of this pathway may lead to tumorigenesis.

Table S1. Characteristics of cell lines used in this study

Table S2. Sequences of small interfering (si) RNAs and PCR primers

Table S3. Reagents and resources.

Table S4. Summary of statistical analyses for individual figures

Dataset S1. Identification of HERV-K loci expressed in iPSC

Dataset S2. Proteins co-immunoprecipitated with HERV-K Env

Dataset S3. Genes regulated by HERV-K Env siRNA treatment in iPSC

Supporting Video: Treatment of antibody against HERV-K Env resulted in the detachment of iPSC colony in neural induction medium.

Table S1. Characteristics of cell lines used in this study

iPSC lines used for RNA-Seq study

Name	Origin	Reprogramming method	Gender	Passage used	Reference
NCRM1	CD34+ cord blood cells	Episomal	Male	18	Efthymiou et al., 2014
NCRM5	CD34+ cord blood cells	Episomal	Male	26	Efthymiou et al., 2014
ND2.0	Skin fibroblast	Episomal	Male	31	Chen et al., 2011

iPSC lines differentiated to neurons for Western-blot assay, immunostaining and RT-PCR.

Name	Origin	Reprogramming method	Gender	Passage used
NC 4	CD34+ cord blood	Episomal	Male	16
WTC11	Skin fibroblast	Episomal	Male	35
ND2.0	Skin fibroblast	Episomal	Male	24
PAU	CD34+ adult blood cells	Sendai virus	Female	
ALEX	CD34+ adult blood cells	Sendai virus	Male	

NSC lines used in this study

Name	Gender	Passage used	Used for RNA seq	Differentiated to neurons for RNA seq
NCRM1	Male	7	X	X
BC1	Male	6	X	X
H14	Male	6	X	X
XCL1	Male	6	X	

iPSC lines used for neuronal differentiation for HERV-K LTR methylation study.

Name	Origin	Reprogramming method	Gender	Passage used
505	CD34+ adult blood	Sendai virus	Male	5

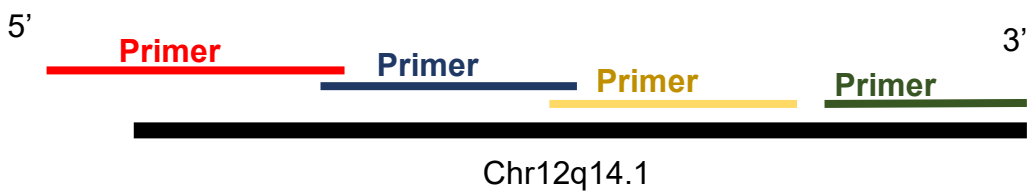
506	CD34+ adult blood cells	Sendai virus	Male	5
516	CD34+ adult blood cells	Sendai virus	Female	5

Table S2a. Sequences of small interfering (si) RNAs

Name	Target	Sequences
Gi	HERVK-gag	Sense5' UCCCAGUAACGUUAGAACCUU 3' Antisense3' UUAGGGUCAUUGCAAUCUUGG 5'
Ei 1	HERVK-env	Sense5' CUGACGCAGUUAGCUACAAUU 3' Antisense3' UUGACUGCGUCAAUCGAUGUU 5'
Ei 2		Sense5' GAUUCACUUAUCACAUGGUUU 3' Antisense3' UUCUAAGUGAAUAGUGUACCA 5'
Ei 3		Sense5' GGUUAAGACCAUUGGAAGUUU 3' Antisense3' UUCCAAUUCUGGUAACCUUCA 5'
Ei 4		Sense5' CAUCAAAAGCCCAUUUAAAUU 3' Antisense3' UUGUAGUUUUCGGGUAAAUUU 5"
Ei 5		Sense5' AGAACAUCGUUUCCAGUUAUU 3' Antisense3' UUUCUUGUAGCAAAGGUCAAU 5'

S2b. Sequences and LTR alignment of PCR primers

Primer	LTR alignment to chr12q14.1 LTR	Primers for Bisulphite treated DNA
12q14.1_1F	5' of LTR beginning in LINE2b sequence	AAAGTTGAAAGGAGTAGTAA
12q14.1_1R	141-165	CACAACACATATTTCAAAAAACAC
12q14.1_2F	141-165	GTGTTTTTTGAAATATGTGTTGTG
12q14.1_2R	446-462	AAACATTCCTTCCTCTTTTA
12q14.1_3F	441-462	ATTAGTAAAAGAGGAAGGAATGT
12q14.1_3R	700-723	CAACAAACAAACATATAAACAAA
12q14.1_4F	745-773	GTGATTTTGATATATTTTTTTTTTYGAG
12q14.1_4R	941-956	AATAAATTACCCCTACACACCT
Primer	**LTR alignment to chr19q11 LTR**	
19q11_1F	199-220	TAGAGTTAAATGGATTAAGGGT
19q11_1R	413-436	AACTAAAAAACRATCAAATCTTTC
19q11_2F	468-489	ATTAGTAAAAGAGGAAGGAATGT
19q11_2R	748-776	CACAAAACAATTATAAAAAAAAAATCAAC
19q11_3F	827-846	ATTAAGGGAATTTAGAGGTT
19q11_3R	956-971	ACTAATCAAAAACCTTTTTC



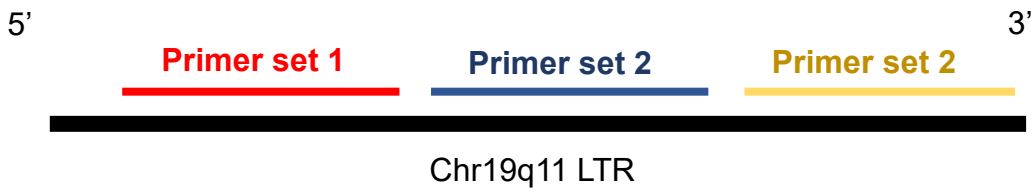


Table S3. Reagents and resources

Reagent or Resource	Source	Identifier
Antibodies		
mouse monoclonal anti-nestin	Abcam	Cat# ab22035, RRID:AB_446723, Lot# GR312885-12
mouse monoclonal anti-β-III-tubulin	Promega	Cat# G7121, RRID:AB_430874, Lot# 0000206749
rabbit anti-MAP2	Abcam	Cat# ab32454, RRID:AB_776174, Lot# GR220420-1
mouse anti-Human TRA-1-60 monoclonal IgM Antibody	Abcam	Cat# ab16288, RRID:AB_778563, Lot# GR160693-1
rat anti-SSEA-3 antibody	Millipore	Cat# MAB4303,RRID:AB_177628, Lot# NG1869267
mouse monoclonal anti-CD98	Abcam	Cat# ab23495, RRID:AB_447477, Lot#GR3180054-3
rabbit anti-LPCAT1	Proteintech Group	Cat# 16112-1-AP, RRID:AB_2135554, Lot# 00007409
mouse monoclonal anti-HERV-K Env	Austral Biologicals	Cat# HERM-1811-5, RRID:AB_10891119, Lot# MA180709HK4
mouse monoclonal anti-HERV-K Env	Geneuro	ID# K06 GN
rabbit anti-HERV-K Env	GeneScript	Custom made
rabbit anti-HERV-K Gag	GeneScript	Custom made
rabbit anti-S6 ribosomal protein	Cell Signaling Technology	Cat# 2217, RRID:AB_331355, Lot# 7
rabbit anti-phospho-S6 ribosomal protein ((Ser235/236)	Cell Signaling Technology	Cat# 4858, RRID:AB_916156, Lot# 16
mouse monoclonal anti-puromycin	Millipore	Cat# MABE343, RRID:AB_2566826, Lot# 3045550
anti-Flag mAb	Sigma	Cat# A8592, RRID: AB_439702, Lot# SLBV3799
anti-mouse Alexa Fluor 488	Thermo Fisher	Cat# A-11001, RRID:AB_2534069, Lot# 1484573
anti-rabbit Alexa Fluor 594	Thermo Fisher	Cat# A-11012, RRID:AB_2534079,

		Lot# 1596064
anti-mouse IgM Alexa Flour 488	Thermo Fisher	Cat# A21042, RRID:AB_141357, Lot# 1387726
anti-mouse IgG Alexa Flour 546	Thermo Fisher	Cat# A-11003, RRID:AB_2534071, Lot# 1885911
Mouse IgG2a control antibody	Thermo Fisher	Cat# MA1-10418, RRID:AB_2536786, Lot# QF2043478
Monoclonal mouse-anti-β actin	Sigma	Cat# A5441, RRID:AB_476744,
Monoclonal ouse anti-beta Tubulin	Millipore	Cat# 05-661, RRID: AB_309885, Lot# 2207268
Peroxidase IgG Fraction Monoclonal Mouse Anti-Rabbit IgG	Jackson Immuno Research Lab	Cat# 211-032-171, RRID:AB_2339149, Lot# 115949
Cells		
WA09-FI	WiCell	hPSC Reg ID: WAe009-A
Mouse (ICR) inactivated embryonic fibroblasts	Thermo Fisher	Cat# A24903
HEK 293		
Kits		
hES/iPS cell characterization kit	Applied Stemcell	Cat# ASK-3006
RNeasy Mini Kit	Qiagen	Cat# 74104
total RNA purification kit	Norgen	Cat# 17200
QuantiTect reverse Transcription Kit	Qiagen	Cat# 205311
Q5 High-Fidelity 2X Master Mix	New England Biolabs	Cat# M0492S
RNase-free DNase set	Qiagen	Cat# 79254
one step SensiFAST SYBR kit	Bioline	Cat# BIO-84005
TOPO TA cloning kit	Thermo Fisher	Cat# K200040
DNA purification kit	Clontech	Cat# 746986.20
Global DNA Methylation LINE-1 kit	Active Motif	Cat# 55017

CellQuanti-blue cell viability assay kit	BioAssay Systems	Cat# CQBL-05K
TransIT-siQUEST® Transfection Reagent	Mirus	Cat# MIR2114
Pierce RIPA buffer	Thermo Fisher	Cat# 89901
CytoTune-iPS Reprogramming kit	Thermo Fisher	Cat# A13780-01
Pierce BCA protein assay kit	Thermo Fisher	Cat# PD 202746
Lipofectamine 3000	Thermo Fisher	Cat# L3000-008
Dynabeads Protein G	Thermo Fisher	Cat# 10004D
P3 Primary Cell 4D-Nucleofector® X kit (pmaxGFP plasmid)	Lonza	Cat# PBP3-02250
Amersham ECL™ Western blotting detection reagents	GE Healthcare	Cat# RPN2106
EasySep human progenitor cell enrichment kit	StemCell Technologies	Cat# 19056
Media and Supplements		
Essential 8 Medium	Thermo Fisher	Cat# A1517001
PSC Neural induction medium	Thermo Fisher	Cat# A1647801
StemSpan serum free expansion medium	StemCell Technologies	Cat# 09600
IMDM	Thermo Fisher	Cat# 12440061
NeuroBasal	Thermo Fisher	Cat# 21103049
DMEM/F12	Thermo Fisher	Cat# 10565018
B27 (+vitamin A)	Thermo Fisher	Cat# 17504044
N2	Thermo Fisher	Cat# 1780240
NEAA	ThermoFisher	Cat# 11140-50
GlutaMax	Thermo Fisher	Cat# 35050061
Antibiotic antimycotic	Thermo Fisher	Cat# 15240062
Recombinant human BDNF	Peprotech	Cat# 450-02
Recombinant human GDNF	Peprotech	Cat# 450-10
Recombinant human TPO	Peprotech	Cat# 300-18

Recombinant human Flt-3 ligand	Peprotech	Cat# 300-19
Recombinant human SCF	Peprotech	Cat# 300-07
Recombinant human IL-6	Peprotech	Cat# 200-06
Recombinant human IL-7	Peprotech	Cat# 200-07
Matrigel (Growth factor reduced)	Corning	Cat# 354230
PDL	Sigma	Cat# P6407-5MG
Accutase cell dissociation solution	Thermo Fisher	Cat# A11105-01
Cyclic AMP	Sigma	Cat# A9501
Fetal bovine serum	Gemini	Cat#100-106
L-Ascorbic acid	Sigma	Cat# A92902

Chemicals and solutions		
DAPI	Roche	Cat# 10236276001
DRAQ5	Thermo	Cat# 62251
EDTA	Mediatech	Cat# 46-034-CL
DPBS	Thermo Fisher	Cat# 14190-136
Saponin	Sigma	Cat# 47036
paraformaldehyde	EMS	Cat# 15710
Triton X-100	AmericanBio	Cat# AB02025-00100
Tween-20	Sigma	Cat# P2287-500ML
Glycerol	Fisher Bioreagents	Cat# BP229-1
Goat serum	Sigma	Cat# G9023-10ML
Non-fat milk	BioWorld	Cat# 30620074-2
BSA	Sigma	Cat# A2153-50G
complete proteinase inhibitor cocktail tablets	Roche	Cat# 11873580001
phosphatase inhibitor cocktail 3	Sigma	Cat# P0044
4-10% Tris-glycine polyacrylamide gels	Invitrogen	Cat# NP0335

SizeSelect™ Agarose Gels	Invitrogen	Cat# G661012
Rapamycin	Sigma	Cat# R8781
Puromycin	Sigma	Cat# P8833
Services	**Company or Institution**	**Location**
iPSC karyotyping	WiCell	Madison, WI
RNA-Seq	New York Genome Center	New York, NY
iPSC generation	NHLBI iPSC Core Facility	Bethesda, MD
FACS	NIH FACS Core Facility	Bethesda, MD
PCR product sequencing	Genewiz	Frederick, MD
	Eurofins Genomics	Louisville, KY

Table S4.Summary of statistical analyses

Figure number	Test used	n	Descriptive stats	P value	Degrees of freedom & F/t value	Post hoc multiple comparisons test
1a	One sample two tailed t-test	4	Error bars are mean ± SEM	P=0.0009	t(3)=13.18	
1b	One sample two tailed t-test	4	Error bars are mean ± SEM	P=0.0001	t(3)=25.51	
1c	One sample two tailed t-test	5	Error bars are mean ± SEM	P =0.0002	t(4)=13.43	
1d bar2	paired two-tailed Student's t-test	4	Error bars are mean ± SEM	P =0.0001	t(3)=25.71	
1d bar3	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P =0.0053	t(2)=13.69	
3a	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P =0.0463	t(2)=4.486	
3b	One way RM ANOVA and Bonferroni's multiple comparison test	3	Error bars are mean ± SEM	P = 0.0319	F(2,2)=9.206	P>0.05
3c	one-way RM ANOVA and Bonferroni's multiple comparison test	3	Error bars are mean ± SEM	P=0.0132	F(2,2)=15.43	Ei vs Nsi <0.05
3e	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P=0.0212	t(2)=6.765	
3f	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P=0.0171	t(2)=7.554	
4b	one-way RM ANOVA and Bonferroni's multiple comparison test	3	Error bars are mean ± SEM	P = 0.0136	F(2,2)=15.15	Vs ctrl: P <0.05 Vs Nsi: P<0.05
5c	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P = 0.0137	t(2)=8.459	

7g	Student's t-test	3	Error bars are mean ± SEM	P=0.0281	t(2)=5.837	
S3a	paired two-tailed Student's t-test	3	Error bars are mean ± SEM	P=0.0123	t(2)=8.931	
S3b	one-way RM ANOVA and Bonferroni's multiple comparison test	3	Error bars are mean ± SEM	P = 0.0002	F(4,2)=6.625	1.5 µM: P <0.05 2 µM: P < 0.05
S3ci	unpaired one-tailed Student's t-test	3	Error bars are mean + SEM	P=0.0352	t(4)=2.451	
S3cii	One way ANOVA and Dunnett's multiple comparison test	3	Error bars are mean + SEM	P=0.0225	F(2,5)=8.899	P<0.05