Supplemental Data

Silencing of Proviral Expression in Embryonic Cells: Efficiency, Stability, and Associated Chromatin Modifications

Sharon Schlesinger and Stephen P. Goff

Supplemental methods

Cell culture and stable RNAi cell line production and transduction

F9, RAT2, NIH 3T3 and 293T cells were cultured in DMEM with 10% FBS, 2 mM glutamine, 1000 U/mL penicillin, 100 mg/mL Streptomycin. E14 ES cell line and primary ES cells taken directly from the embryo ICM were used as ES cells, and were cultured on gelatinized tissue culture plates in ES cell media (Dulbecco modified Eagle medium (DMEM) supplemented with 15% defined fetal bovine serum (HyClone Cat. SH30070), 100 IU/mL penicillin, 100 mg /mL streptomycin, 2 mmol/L L-glutamine, 5 mg /mL MEM non-essential amino acids, 0.12 mmol/L β-mercaptoethanol, and 1000 U/mL leukemia inhibitory factor (ESGRO, Millipore). All cells were cultured at 37°C in 5% CO2. RNAi knockdown was performed using retroviral vectors encoding shRNAs. The VSV-G pseudotyped viruses were generated by cotransfection of 293T cells using Fugene 6 (Roche) with a mixture of pSUPER.Retro.Puro (oligoengine) and the helper plasmids, pCMV-intron and pMD.G, using a 1.5:1:1 ratio of the three plasmids in each transfection. All viruses were harvested 48 h post-transfection, filtered (0.45-mm filter, Pall Acrodisc), aliquoted, and stored at -80°C. Neat medium containing retroviruses (with 8 mg/ml polybrene) was used to infect F9 cells which were then allowed to recover for 48 h followed by selection with 20 mg /ml Puromycin. For viral transduction assays, viruses were prepared as above using either LJ-PAdMLPEnh- or LJB2-ADMLPEnh- vectors [1] or pNCA-GFP vectors [2]. Retroviral stocks were then serially diluted (for titer determination), and added to F9 and RAT2 cells (seeded at [F9] 3.5×10^3 and [RAT2] 2×10^3 cells per cm^2 the day prior to transduction) in the presence of 8 μg of Polybrene/ml. G418-containing selective media was added 48 h post transduction at 1mg/ml for RAT2, and at 0.5mg/ml for F9 cells, colonies were counted after 14 to 18 days of selection. Each experiment was performed in duplicate and repeated three times.

Chromatin Immunoprecipitations (ChIP)

10^7 cells were crosslinked with 1% formaldehyde at room temperature for 10 mins. chromatin was extracted and then sonicated to an average size of 300–3,000 bp. Immunoprecipitation was carried out by using a Upstate ChIP assay or Magna ChIP™ kit as recommended by the manufacturer (Millipore) and then purified using QIAquick PCR purification kit (Qiagen). Antibodies used (about 5 μg per 10–30 μg of DNA) were: Anti-trimethyl-Histone H3 (Lys9) (07-442, Millipore), Anti-trimethyl-Histone H3 (Lys27) (07-449, Millipore) and Anti-acetyl-Histone H3 (06-599, Millipore). IgG antibody was used as negative control (sc-2027, Santa Cruz) Amplification was carried out by real-time PCR, and the bound/input values were then normalized by setting the negative control gene or the IgG results to 1. Multiple assays of the same sample or the same gene sequence were analyzed in separate immunoprecipitations. All immunopercipitations were repeated 3-5 times. Primer sequences used for qPCR are listed in Supplemental Table 1.

Bisulfite analysis

Bisulfite conversion of genomic DNA was carried out using the Zymo EZ DNA Methylation-Gold™ Kit according to the manufacturer's instructions. PCR primers were designed using Methyl Primer Express software version 1 (https://www2.Appliedbiosystems.Com). After PCR amplification, products were extracted from gels using the Qiagen MinElute Gel Extraction Kit, cloned into the Promega pGEM-T vector system, transfected into Promega JM109 competent bacteria, and plated on LB-Agar + Amp + X-gal and IPTG. The DNA from 10–12 white clones of each ligation was subjected to sequencing with T7 or Sp6 primers. Primer sequences are listed in Supplemental Table 1. Sequence results were analyzed with the BiQ analyzer [3], using all standard quality control steps.

ERVs Expression

Total RNA was prepared with the TriPure isolation reagent (Roche), and treated with DNase I (Ambion). For qRT–PCR analysis, cDNA was generated using Super-Script III (Invitrogen) and random hexanucleotide pd(N)6 primers. Control reactions lacking reverse transcriptase were systematically verified for the absence of products. qPCR was performed using SYBR Green master mix (Roche). All reactions were

performed in triplicate. Standard curves were generated for each primer pair (see table S1), and expression levels were normalized relative to three control genes (Cyclophilin-A, UBC and Gapdh). Error bars represent SD. Primer sequences used for RT-qPCR are listed in Supplemental Table 1.

Pluripotency analysis

ES-specific markers were examined using BD Human and Mouse Pluripotent Stem Cell Analysis Kit (560477, BD) followed by multicolor Flow Cytometry (LSR II; BD).

Virus copy number determination

The difference in threshold cycle (CT) values (ΔCT) between Q-PCR assays using two of the primers on the pro-virus, 40NT and PBS, was used to monitor proviral DNA copy number. The Gapdh primer set for the genomic DNA was used to normalize the amount of genomic DNA. A one copy number standard was established by infecting NIH 3T3 cells at a very low multiplicity of infection (MOI) with the MLV-GFP vector and sorting the GFP(+) cell population (10% of the total), ensuring that a single copy of GFP virus was present [4]. The ratio of GFP to genomic DNA (Gapdh) in this sample is normalized to 1 (one copy of provirus per cell genome), and this sample is taken as the calibrator. The differences in ΔCTs (ΔΔCT) for the samples of interest and the calibrator are used to estimate the relative quantity (RQ) of provirus by using the formula $RQ=2^{-\Delta\Delta CT}$. The copy number values given were obtained by averaging results from three PCR reactions. Uninfected NIH 3T3 cells were used as negative control. In the flow analysis results presented, all numbers were normalized to the infection efficiency as seen by the % of NIH3T3 expressing cells. The numbers are given as GFP-positive F9 cells/GFP-positive NIH3T3 cells x100; this is to correct for variation in multiplicity with different viruses.

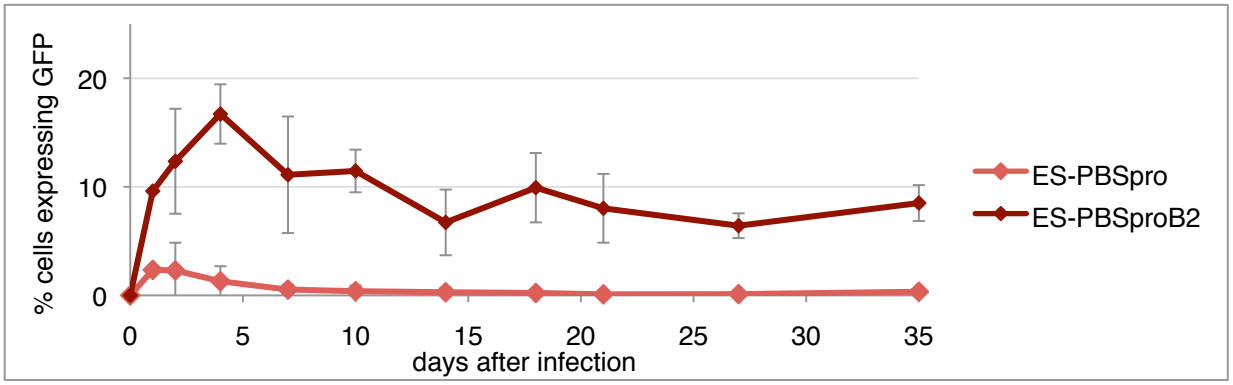
Supplemental references

1. Modin C, Lund AH, Schmitz A, Duch M, Pedersen FS (2000) Alleviation of murine leukemia virus repression in embryonic carcinoma cells by genetically engineered primer binding sites and artificial tRNA primers. *Virology* **278:** 368-379
2. Ooi SK, Wolf D, Hartung O, Agarwal S, Daley GQ, Goff SP, Bestor TH (2010) Dynamic instability of genomic methylation patterns in pluripotent stem cells. *Epigenetics Chromatin* **3:** 17
3. Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T (2005) BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21:** 4067-4068
4. Katz RA, Jack-Scott E, Narezkina A, Palagin I, Boimel P, Kulkosky J, Nicolas E, Greger JG, Skalka AM (2007) High-frequency epigenetic repression and silencing of retroviruses can be antagonized by histone

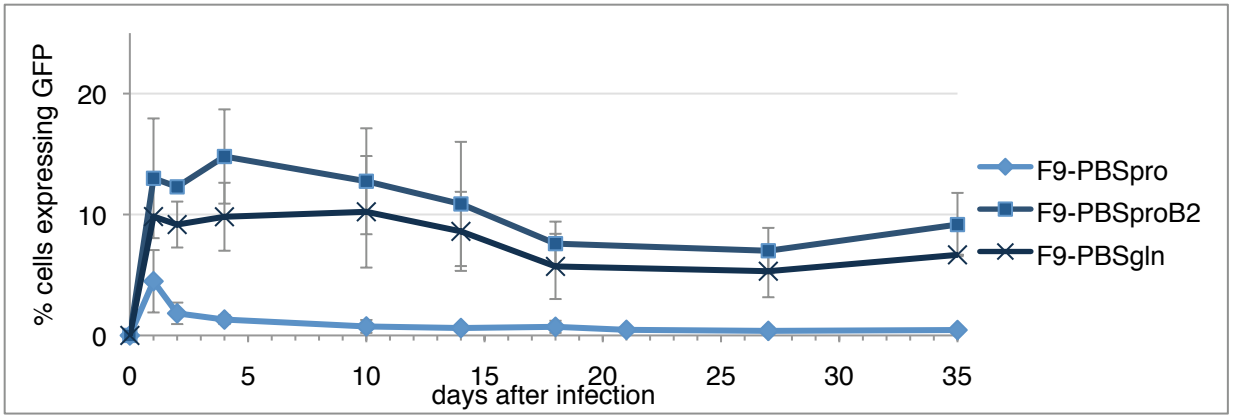
deacetylase inhibitors and transcriptional activators, but uniform reactivation in cell clones is restricted by additional mechanisms. *J Virol* **81:** 2592-2604

Figure S1.

A. Normalized and averaged GFP expression after ES cells infection



B. Normalized and averaged GFP expression after F9 EC cells infection



C. ZFP809 expression

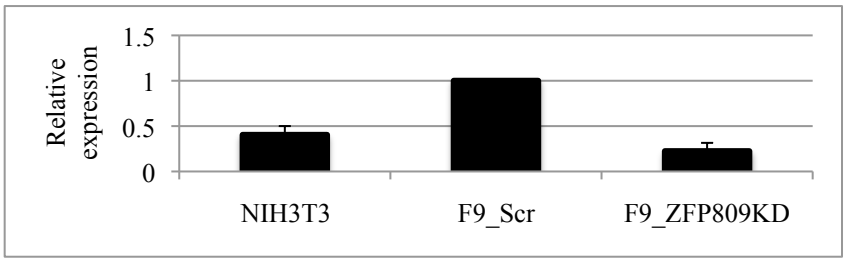
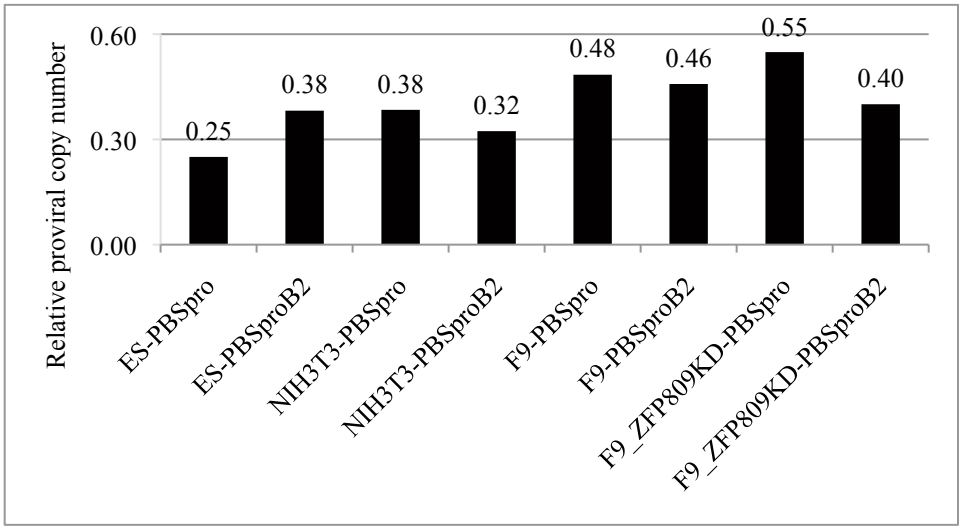


Figure S1.

(A) Flow analysis results at different time points after infection (day 0) of ES cells with wt (pro) and mutant (proB2) PBS virus. N=5. (B) Same assay, on F9 EC cell line with wt and two different mutant PBS viruses. N=5 (PBSpro and PBSproB2) or N=3 (PBSgln). (C) ZFP809 mRNA levels in the indicated cell populations normalized for F9_Scr = 1. Expression levels were determined independently by RT-qPCR analysis. N=4.

Figure S2.

A. Infected cells from Figure 1A, B



B. Infected and sorted cells from Figure 1C, D

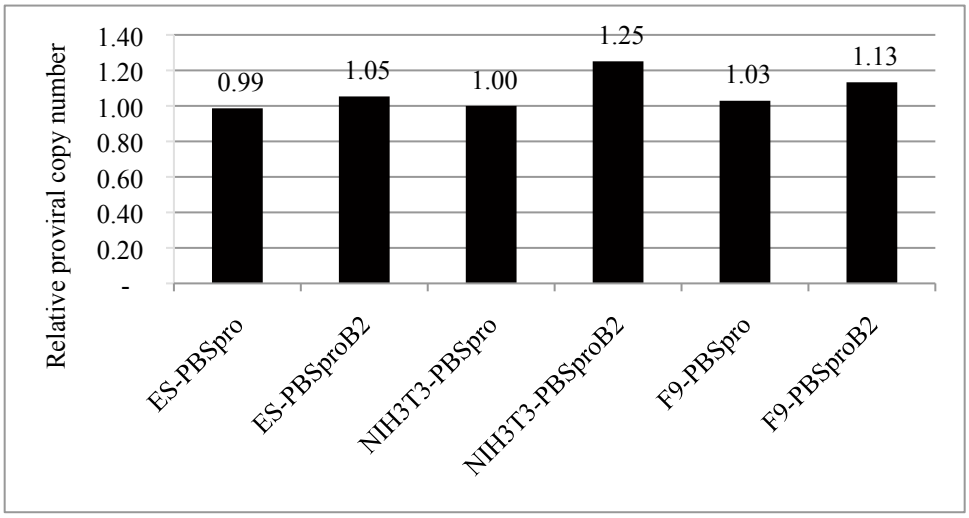


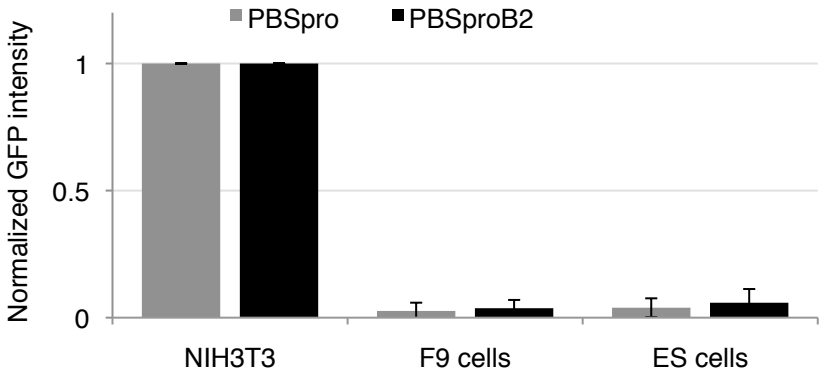
Figure S2. Proviral DNA copy number

One representative experiment was analyzed for infection efficiency by qPCR.

(A) Infected cell lines in Figure 1A,B, (B) infected and sorted cells from figure 1C, D.

Figure S3

A. Mean flow analysis



B. Flow analysis – three examples

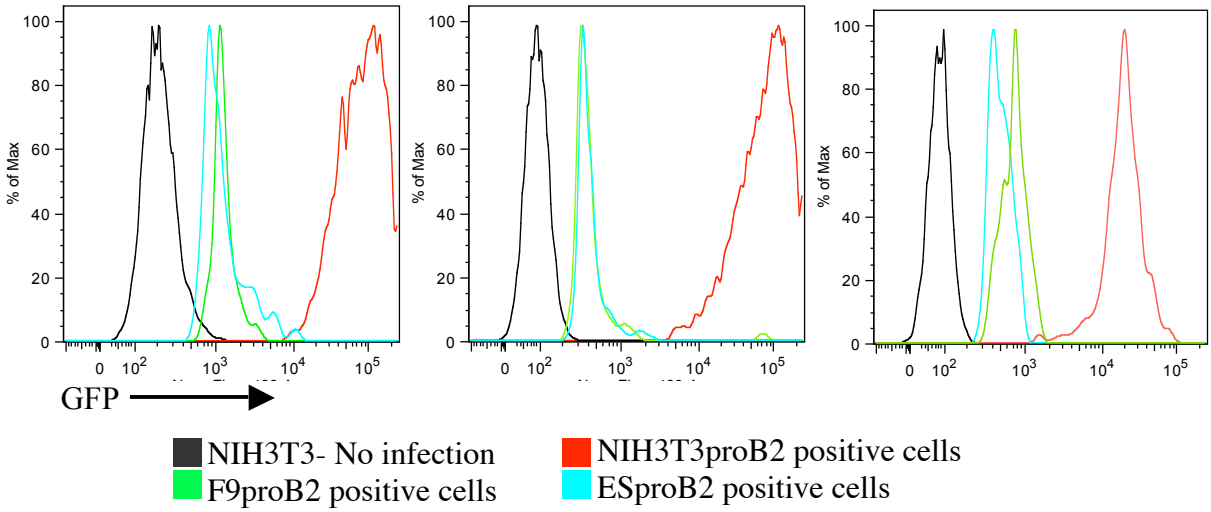
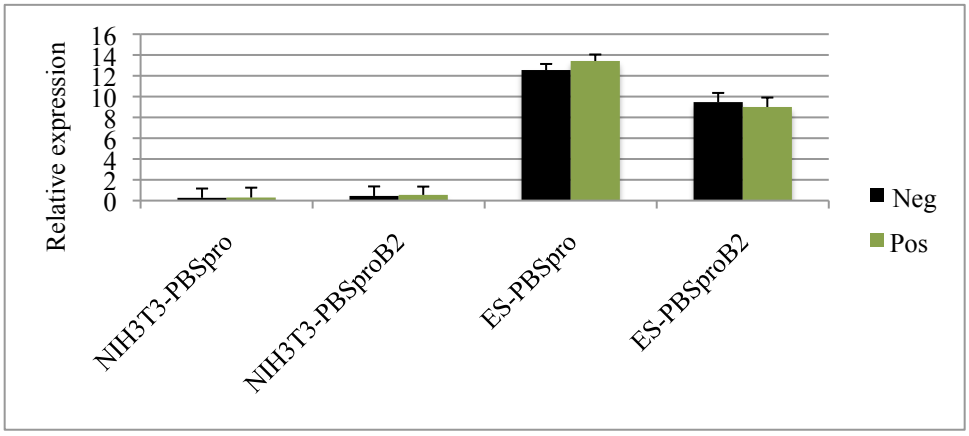


Figure S3.

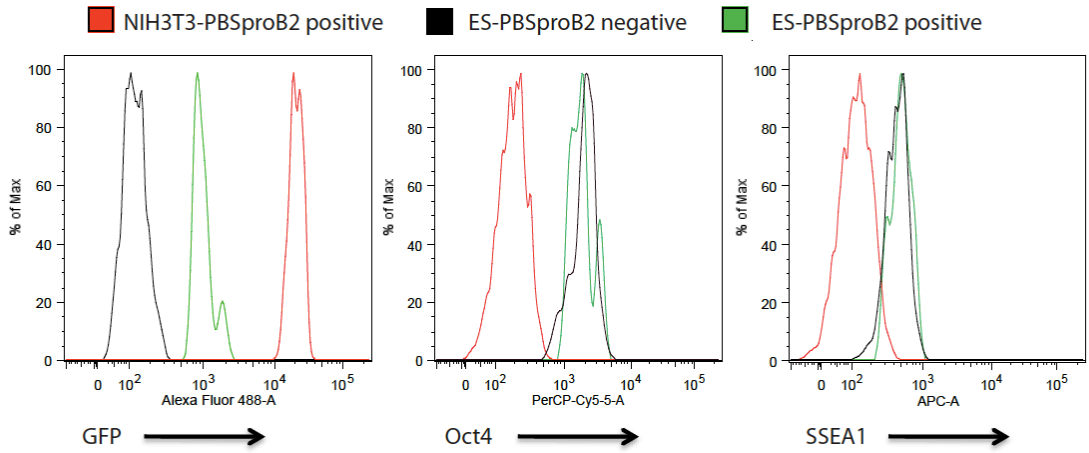
(A) Flow analysis of GFP levels by fluorescence intensity measurement analyzed by flow cytometry. The results demonstrate that fully differentiated cells have significantly higher levels of GFP expression than embryonic cells, N=5. For Copy number analysis see Fig S2A. (B) Raw flow analysis results of three (out of five) cell populations infected and sorted for GFP expression. No infection control and GFP positive sorted cells are shown. Analysis was done immediately after the sorting.

Figure S4

A. Oct4 expression



B. Flow analysis



C. GFP expression

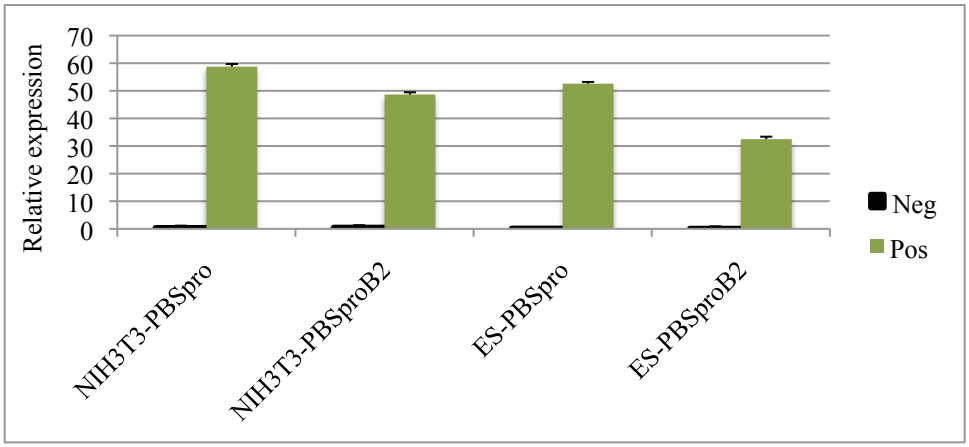
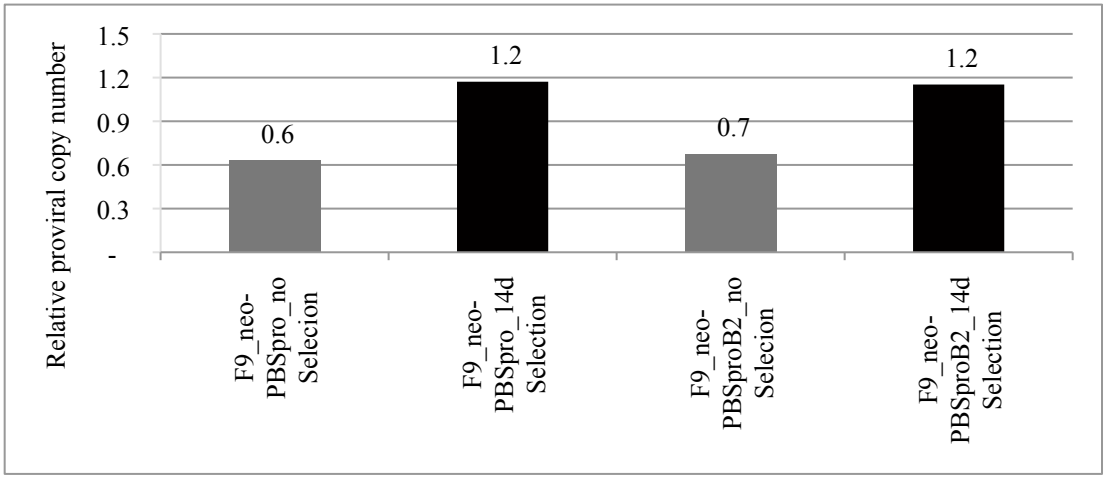


Figure S4. (A) Oct4 mRNA levels in the indicated cell populations sorted for negative (Neg) and positive (Pos) GFP expression. Expression levels were determined independently at the day of sorting by RT-qPCR analysis. (B) Flow analysis of sorted cell populations. Panel I: GFP levels. Panel II: Oct4 levels by antibody stain. Panel III: SSEA1 (stage specific antigen 1) level by antibody stain. Red curves: NIH3T3 cells infected by PBSproB2 virus, sorted for GFP. Black curves: ES cells infected by PBSproB2 virus, sorted for GFP negative; Green curves: ES cells infected by PBSproB2 virus, sorted for GFP positive. The results shows that ES cells positive and negative for GFP (I) do express the ES cell markers Oct4 (II) and SSEA-1 (III), while NIH3T3 cells do not. (C) GFP mRNA levels in the sorted cells.

Figure S5

A. Proviral DNA copy number



B. Oct4 expression

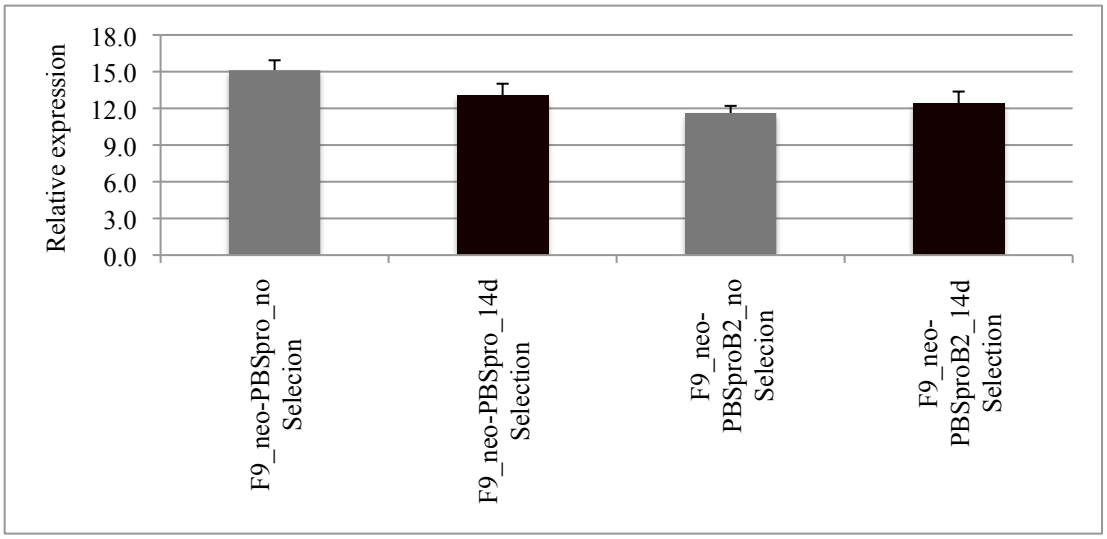
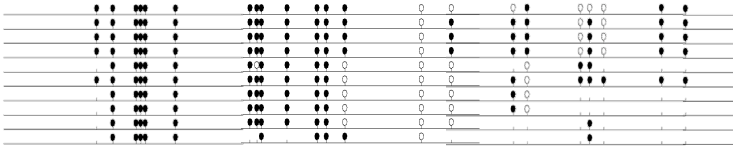


Figure S5. Analysis of proviral DNA copy number and ES cell specific markers in selected and unselected cell populations. (A) DNA copy number in the indicated cell population as determined by qPCR. (B) Oct4 levels as determined by RT-qPCR.

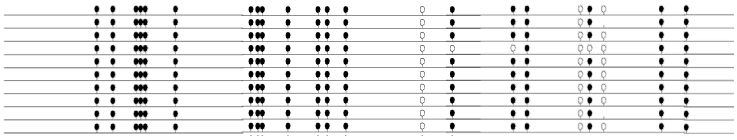
The results show that surviving long term selection is not due to high copy number or cell differentiation.

Figure S6

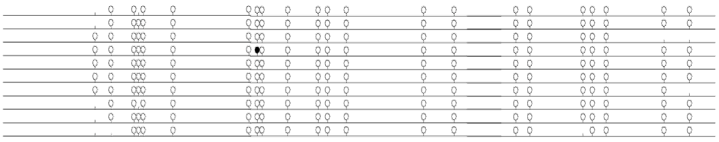
A. ES_ PBSpro - 14 days after infection, 82% methylated



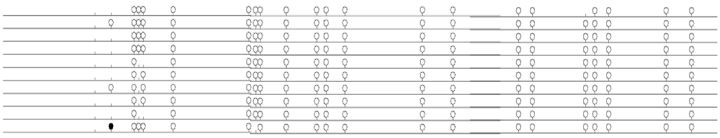
B. ES_ PBSproB2 - 14 days after infection, 89% methylated



C. NIH3T3 PBSpro - 14 days after infection, <1% methylated



D. NIH3T3_ PBSproB2 - 14 days after infection, <1% methylated



E. ES_ Oct4 ctrl

F. NIH3T3 _Oct4 ctrl

Figure S6. De novo methylation of proviral DNA is PBSpro independent. Retrovirus reporter construct with long terminal repeats (LTRs) modified to allow identification of reporter provirus against background of endogenous retroviruses was used to infect wild-type ES cells and methylation status examined (A). Same assay, with PBSproB2 virus infected cells (B). NIH3T3 ware infected with the two viruses and analyzed (C,D). As a control, the methylation status of Oct4 promoter region was analyzed in ES cells (E) and NIH3T3 cells (F). Open and filled circles represent unmethylated or methylated cytosines, respectively.