

Supplementary Figure 1

A

Name	Location	Start position	End position	Annotation	RefSeq accession number
Oct4P1	Chr 3	129433673	129435000	Mus musculus predicted gene 5712 (Gm5712), non-coding RNA	NR_033594.1
Oct4P2	Chr 1	159293805	159294189	N.A.	-
Oct4P3	Chr 14	17105602	17106308	N.A.	-
Oct4P4	Chr X	52878731	52880395	Mus musculus predicted gene 6539 (Gm6539), non-coding RNA	NG_021129.1
Oct4P5	Chr 6	103964077	103964639	N.A.	-

B

Oct4P1 (182 bp)

...GGCGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTGGGAATTCGACGTCGGGATGCTGTGAGCCAAGGCAAGGAAGAGAGGTA GACAAGAGAACCTGAAGTTTTGGGGTTAAATCTTTTACTGAGGAGGGATTAAAGCACAAAGGGGGTGGGGTGGGGATGGGATAGGATAAGAAGTTCAAGTATGCTGTTGATCTGGAGC CTGTCTGTCACTCATCCTTAAATAAAGAAATCCCAAGGGCGAATTCACGACACTGGCGGCCGTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCC TGTGTGAAATTTGATCCGC...

Oct4P2 (394 bp)

...TAGGGCGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTGGGAATTCCTTAAAGAACAAAATGATGAGTGACAGACAGGCCAGGCT CCTGATCAACAGCATCACTGAGCTTTTCCCATCCCACCCCCACCCCTGTTGTGCTTTTAAATCCCTCAGTAAAGAATTTAACCCCAAAGCTCCAGGTTCTTGTCTACCTCCCTT GCCTTGGCTCAGCATCCCCAGGGAGGCGCTGGTGCCTCAGTTTGAATGCATGGGAGAGCCAGAGCAGTACGCGGAAACAGAGGGAAAGGCCTCGCCCTCAGGAAAAGGGACTGAGT AGAGTGTGGTGAAGTGGGGCTTCCATAGCCTGGGGTGCCAAAGTGGGACCTGGGGGAGAGGAAAGGATACAGCCCCCTGTCTTCTACACTGGGTGAAACTTGAGCATACGGAA TCCCAAGGGCGAATTCAGCAGCACTGGCGGCCGTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGAAATTTGATCCGCTCACAAATTC ACACAACATA...

Oct4P3 (214 bp)

...AGGGCGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTGTCACTACCTGGCTCTCCATGCATTGAGACTGAGGCACCCAGCAC TCCCTGGGGATGATGTGAACAAAGGCAAGGAAGGGAGGTAGACAAGAGAACCCGGAGCTTTGGGGTTAAATCTTTTATTGAGGAGGGATTAAGCAACAAGGGGGTGGGCATGGGATA AGAAGTCCAGTATGCTGTTGATCAGGAGCTCGTCCACTCATCAAGGGCGAATTCAGCAGCACTGGCGGCCGTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATG GCATAGCTGTTTCTGTGT...

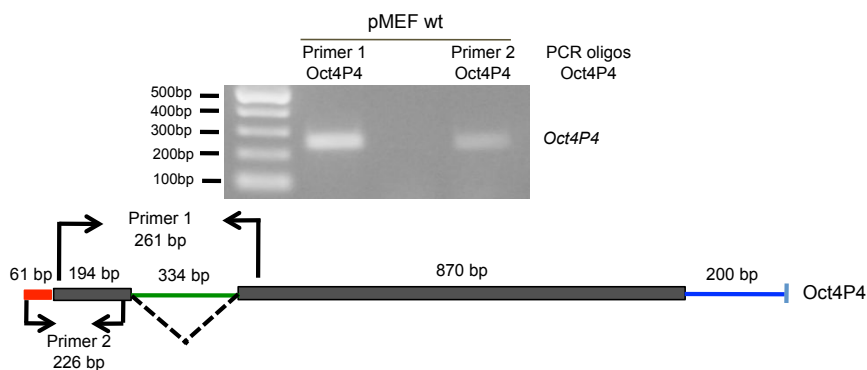
Oct4P4 (160 bp)

...CGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTGTGGCACCTGGCTTTAGACTTTATCTTACTCCCATAGGGAGTGGGGAGG ATGGGTAAGGAAAGCTGGAAGCTGGGCTGGGTGGACCCTTTGACCTGGATAAGCTCCCAAGGGCCCCAAGTGGGCTGGAATGGGGCCAGAGGGTGGACAAGGCTTGAAGGGCGAA TTCCAGCAGCACTGGCGGCCGTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTG...

Oct4P5 (155 bp)

...AGGGCGATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTAGCTCCCATAGCCTGGGGTATCAAAGTGTCAAAGTGGGACCTGGG GGCAGAGGAAAGGATACAGCCCCCTTCCCCCGCAAGGCGTCCCTGTGACCTCATACTCCTCTCTGTGGGAATACACAATGTTTGATCGTTGGCCCTTCTGGCACCGAAGGGCGAAT CCAGCAGCACTGGCGGCCGTACTAGTGGATCCGAGCTCGGTACCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTG...

C



D

Oct4P1
MetAGHLASDFAFSPHQVGA**Met**GQQGWSRAGWTL**EPG** **Stop**ASKGLWVVGLE**Stop**DQA**Stop**RYWDLP**Met**STGI**Stop**VLRRDGILWTSWVRGD
 FAA**Stop**EPGKSRGEGQLRGLSLLWALCRPQC**REV**GD**EG**TTSRGVPGHESPAEGAREVCAAEAEEDHLGVHPDRRGAPGRSLQKGV
 QTTIYCFEARQLSL**Met**TT**Met**HKLRPLLEKWVEEAEDNENLQ**IE**CK**SET**MetVQAWKKKSASIEHVRQSLNEFLK**CE**PEPS**PO**QIT**S**ISIT**Met**
 QLG**Met**EKD**Met**VRGWFCNRRQK**GR**SS**IE**YSQREKYEA**FR**GC**LS**GRGVGLY**PF**LC**Q**APL**WH**PKL**WE**PP**LH**HT**TR**YLS **Stop**GRGLSLCP
 LSLWALPCIQETPALPGDAVSGQKEER**Stop**TREPEVLGLNSFT**EE**GLKARRGWGGDGIG**Stop**EVQ**Stop**CC**Stop**SGACLSLILK**Stop**RLG
 HTVDS

Oct4P2
 QGGLYPFPCPVPTLAPQA**Met**EAP**TS**PH**ST**Q**S**L**FL**RAR**PF**LP**FS**LLWALPCIQTEAPALPGDAVSGQK**GR**QENLELWG**Stop**ILLRRD
StopKHNRGGGWGGERSSV**Met**LLIRSLACL**S**LILFLNKD**WT**Q**Stop**

Oct4P3
 GVQPDNHLSL**Stop**GPAACVSCG**PF**StopRSWWRKP**PT**R**RE**PLGN**Met**OIGGGPAGW**EE**NSSSIENVRW**SL**ENVFLK**CP**KSS**LQ**OIT**S**IAK
 QLGLVKD**V**VRV**W**FCN**WH**QNG**K**RS**SS**TEYS**Q**PEEF**IE**IQGHLYWGVGV**AV**SE**FL**PQVPTLAPQA**Met**GD**PT**SPH**ST**Q**S**L**FL**WAR**PF**LP
 LSLP**W**LSHAFRLRHQ**HS**LG**Met** **Met** **Stop**TKARKGGRQEN**PE**LWG**Stop**ILLRRD**Stop**KHNRGGHGIRSPV**Met**LLIRSS**S**ITHHFYS**Stop**IKT

Oct4P4
MetA**WH**LALDFIFLLP**Stop**GVGR**Met**G**K**ES**W**SWAG**WT**L**Stop**PG**Stop**APKGPQV**GL**EWGQRVDKAWG**CV**WSPH**IP**QSS**Met**EE**E** **Stop**HIVNL
 RLDWDKCPK**L**AWRLC**S**LRPSWS**Met**S**G**EEL**RE**GLS**Stop**ALYS**PF**ECCEVGEAGTKSPGVSGHESPADRTRVVCQTAAEA**EE**DQ**FG**VQ**PG**Q
 HGAHPGCSLWRSVQPDHHLQL**Stop**GPA**Q**LQELV**Stop**TAAPAGEVGGSSG**Q**Q**Stop**GCSR**Met**Y**CK**P**W**CR**P**ARE**S**R**Stop**ALRTLRRGT**Stop**
 KTHFCSGGSLPCSRSAASSNSLGWR**Met**W**P**CG**S**VT**CT**RRV**NK** **Stop**SSIDYSQVRS**CR**ASFGGD**LY**PF**LC**FQGLILVPRG**Met**G**T**PIS**P**
 H**CT**HW**S**L**F**AR**I**R**P**L**FL**LL**Stop**ALPCIQAEATSLPGGEVSLG**EE**GG**Stop**REP**G**ATTAL**EL**SS**L**TEER**F**KAQ**K**G**Stop**RGR**Stop**SYAVE**CE**L
 SLTILFNLKSLGH

Oct4P5
 T**Stop**KRK**WT**STEN**CL**R**Stop**SL**EN** **Met**F**Q**K**R**K**P**SL**Q**E**IT**S**I**IT**K**QL**R**LE**K**D**V**AR**V**W**F**CT**R**C**Q**K**Q**RS**N**I**V**S**Q**RE**E**Y**VT**G**T**PC**G**GG**GL**
PFLC**Q**V**PT**LL**Stop**Y**PR**L**W**EL**PL**H**HT**LL**SP**FS**Stop**G**Stop**GL**PF**V**PH**CY**GA**LP**CI**QA**Stop**APAL**PG**SD**V**SLG**K**GR**Stop**TK**RT**W**SS**GV**TF**
 Y**Stop**TRIK**STR**L**G**GR**G**IR**SS**V**Met**

E

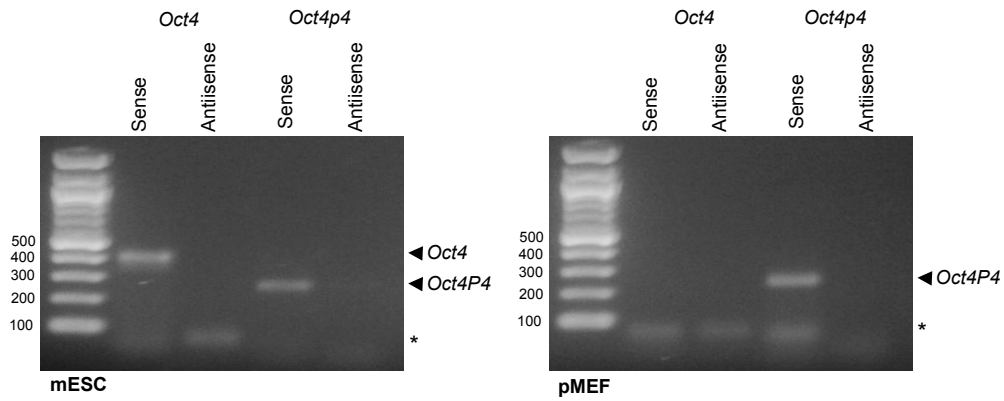
	<i>Oct4P4</i> (molecules/ μ g of RNA)	Absolute Fold (ES set as 1)	Relative Fold (pMEF RNA/ cells= 4.8x ES RNA per cell) ES set 1
ES cells	6882.29	1	1
pMEF cells	906741.14	131.74	632

	<i>Oct4</i> (molecules/ μ g of RNA)	Absolute fold (ES set as 1)	Relative fold (pMEF RNA/ cells=4x ES RNA per cell)
ES cells	51617142.8	-	-
pMEF cells	N.D.	-	-

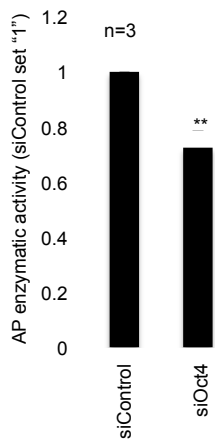
(A) Chromosomal localization, genomic coordinates (UCSC genome browser), annotation and RefSeq accession numbers of *Oct4* pseudogenes describe in this study. Only *Oct4P1* and *Oct4P4* result currently annotated. (B) Sequencing validation of *Oct4* pseudogenes transcripts addressed in this study. Individual *Oct4* pseudogene cDNA templates were specifically amplified by RT-PCR and subcloned before sequencing. (C) Analysis of *Oct4P4* transcripts in mouse embryonic fibroblast (MEF) by semi-quantitative RT-PCR using specific primers identifying spliced and non-spliced *Oct4P4* mRNA. Top panel: *Oct4P4* RT-PCR products visualized on an 1.2% agarose gel. Bottom panel: schematic representation of possible PCR products. Primer pair 1 gives a 261bp product indicative for *Oct4P4* pr-mRNA splicing. Primer pair 2 amplifies a 226bp region upstream the spliced insert, demonstrating that *Oct4P4* transcription is initiated upstream of the spliced insert. (D) Potential peptide sequences encoded by *Oct4* pseudogenes. Potential start codons (Met) are indicated. Stop codons (Stop) are indicated. The maximum ORF predicted for *Oct4P4* spans 27 amino acids. (E) Estimated number of *Oct4* and *Oct4P4* molecules per microgram of total RNA in mESC and pMEFs. *Oct4* was not detectable in pMEF. Absolute fold: difference of total *Oct4P4* (upper panel) or *Oct4* (bottom panel) mRNA levels per μ g of total RNA in pMEFs when compared to mESCs (set as 1). Relative fold: pMEFs (4.8 pg RNA/cell) contain 4.8 fold more total RNA than mESCs (1pg RNA/cell). Indicated values correct for differences in total RNA in mESCs compared to pMEFs. pMEFs contain 632 fold more *Oct4P4* than mESCs.

Supplementary Figure 2.

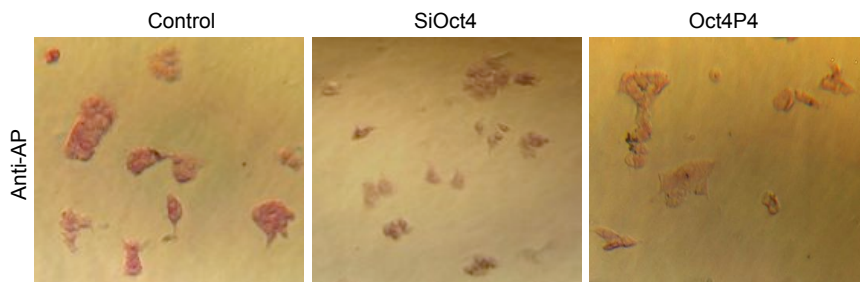
A



B

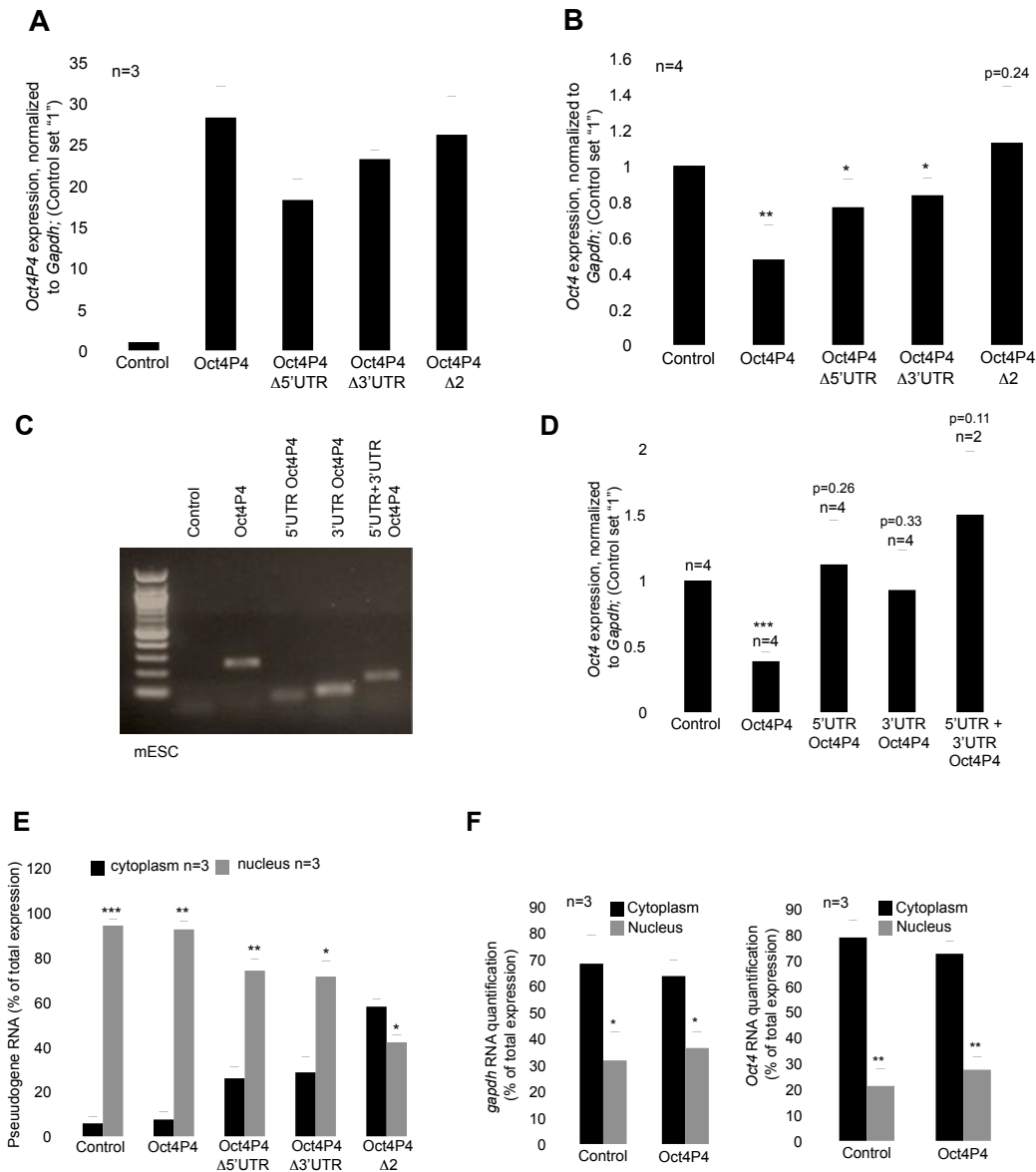


C



(A) Strand specific *Oct4* and *Oct4P4* PCR products obtained from mESC and pMEF cDNA, run on a 1.2% agarose gel. Asterisk indicates primer dimers. *Oct4* was not detectable in pMEF. (B) Alkaline phosphatase (AP) activity in mESCs after RNAi mediated depletion of *Oct4*. (C) AP staining of mESC colonies transfected with the indicated siRNA or expression vector. siOct4 and *Oct4P4* overexpression reduce AP activity.

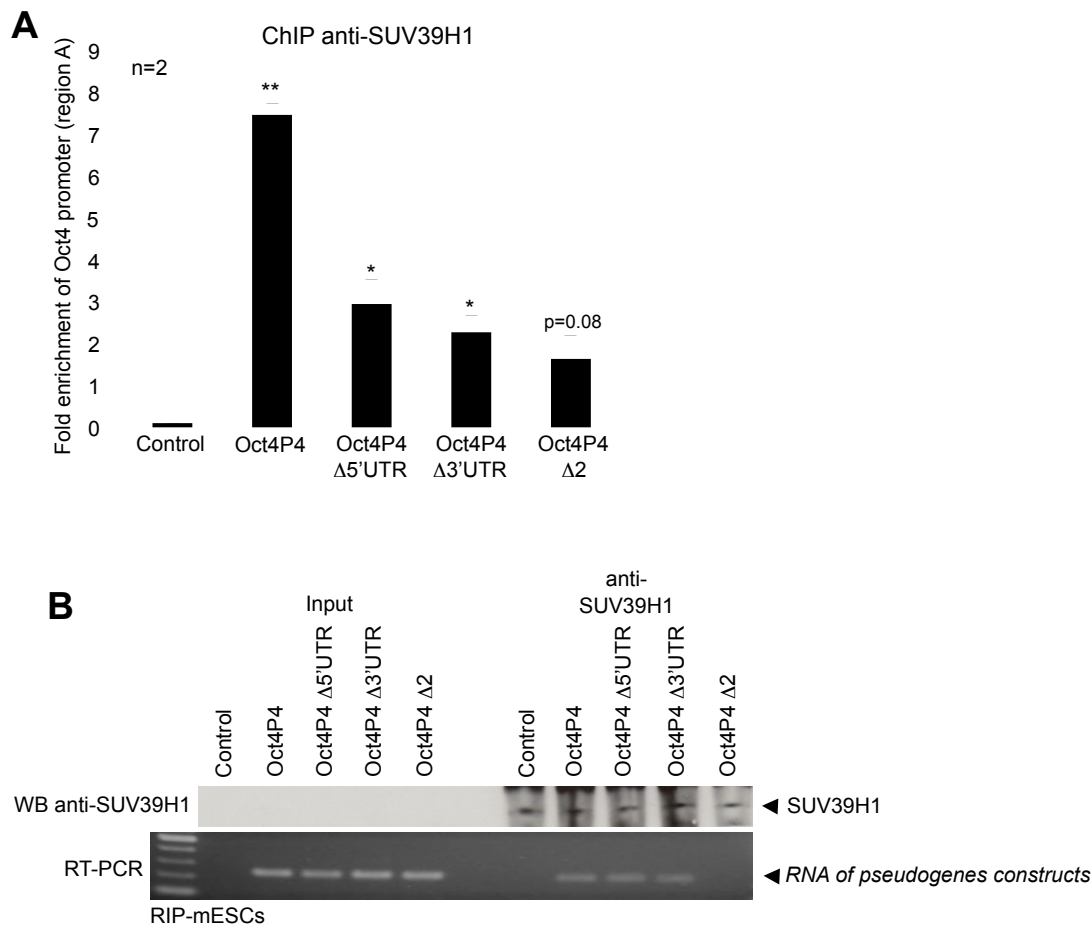
Supplementary Figure 3.



(A) Quantitative RT-PCR analysis using mESCs stably expressing Oct4P4, Oct4P4Δ5'UTR, Oct4P4Δ3'UTR or Oct4P4Δ2. Expression values were normalized against *Gapdh*. (B) Quantitative RT-PCR analysis of ancestral *Oct4* using mESCs stably expressing Oct4P4, Oct4P4Δ5'UTR, Oct4P4Δ3'UTR and Oct4P4Δ2. *Oct4* expression values were normalized against *gapdh*. Ectopic Oct4P4Δ5'UTR or Oct4P4Δ3'UTR reduce ancestral *Oct4* expression less efficiently than full length *Oct4P4*. Ectopic Oct4P4Δ2 does not impact on *Oct4* expression. (C) Oct4P4, Oct4P4-5'UTR, Oct4P4-3'UTR or Oct4P4-5'UTR+3'UTR (fusion of Oct4P4-5' and 3'UTRs) specific RT-PCR products were run on a 1.2% agarose gel. (D) Quantitative RT-PCR analysis using mESCs stably expressing Oct4P4, Oct4P4-5'UTR, Oct4P4-3'UTR or Oct4P4-5'UTR+3'UTR (fusion of Oct4P4-5' and 3'UTRs). *Oct4* expression values were normalized against *gapdh*. Expressing exclusively the 5' and or the 3' UTR of *Oct4P4*, does not result in repression of *Oct4*. (E) Quantitative real-time PCR demonstrating subcellular localization of Oct4P4Δ5'UTR, Oct4P4Δ3'UTR

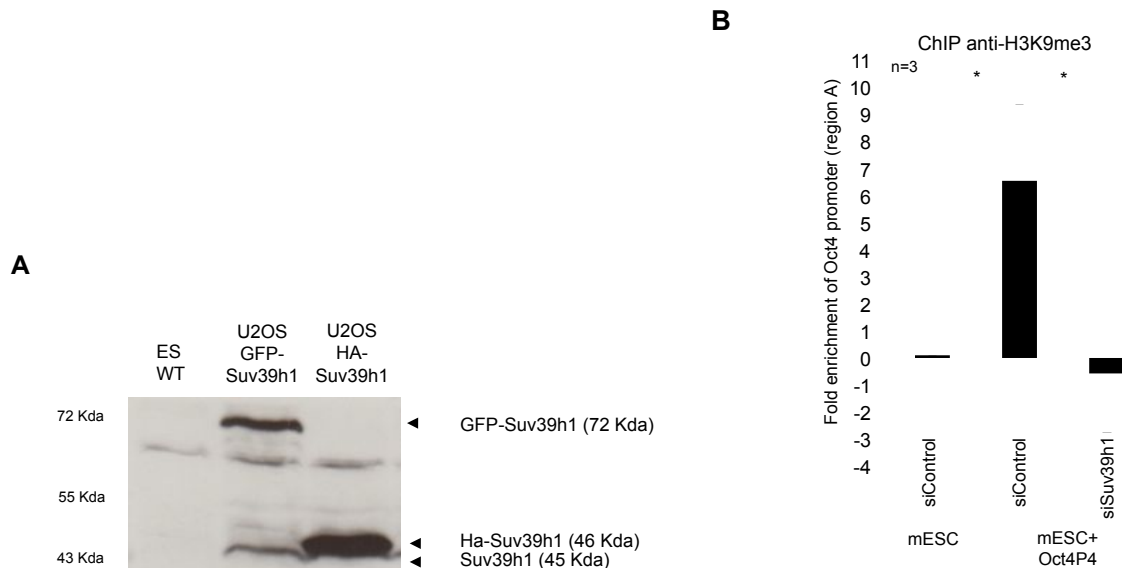
or Oct4P4 Δ 2 in mESCs. Expression values are shown as percentage of total expression. Ectopic Oct4P4 Δ 5'UTR or Oct4P4 Δ 3'UTR show reduced nuclear localization when compared to *Oct4P4*. Ectopic Oct4P4 Δ 2 shows almost equal distribution between nucleus and cytoplasm. **(F)** Quantitative real-time PCR indicating subcellular localization of *gapdh* and *Oct4* in mESCs stably overexpressing *Oct4P4* or a control construct. Expression values are shown as percentage of total expression. Ectopic *Oct4P4* expression does not cause changes of *gapdh* or *Oct4* mRNA localization. . n, refers to the number of independent experiments carried out; error bars indicate standard deviation; a Student's t-test was used to calculate statistical significance values: *, p < .05; **, p < .01; ***, p < .001.

Supplementary Figure 4.



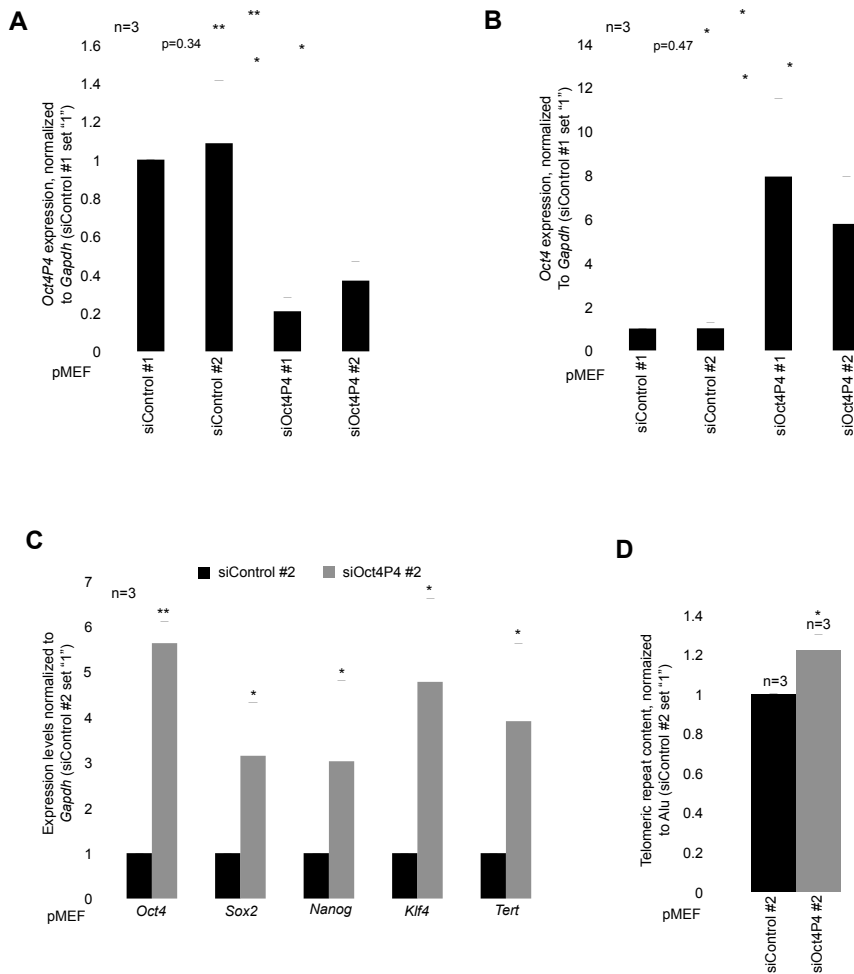
(A) *Oct4* promoter ChIP analysis using SUV39H1 specific antibodies using mESCs transiently transfected with the indicated constructs. Upon deletion of the 5' or the 3' UTR of *Oct4P4*, binding of SUV39H1 to the *Oct4* promoter is reduced. Oct4P4Δ2 does not interact with SUV39H1. **(B)** Anti-SUV39H1 RIP using mESCs overexpressing Oct4P4, Oct4P4Δ5'UTR, Oct4P4Δ3'UTR or Oct4P4Δ2 mESCs. Wild type mESCs were used as control. Immunoprecipitation of SUV39H1 was validated by western blotting (top). Agarose gel electrophoresis after semiquantitative PCR demonstrates the presence of RNA of pseudogenes in anti-Suv39h1 RIP experiments (bottom). Deletion of *Oct4P4* 5'UTR and 3'UTR homology renders *Oct4P4* unable to interact with SUV39H1. Endogenous *Oct4P4* (control) expression levels are too low in mESC to permit *Oct4P4* detection in SUV39H1 immunoprecipitations. n, refers to the number of independent experiments carried out; error bars indicate standard deviation; a Student's t-test was used to calculate statistical significance values: *, $p < .05$; **, $p < .01$.

Supplementary Figure 5.



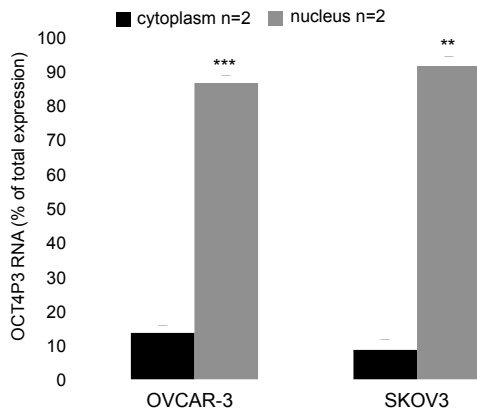
(A) SUV39H1 antibody specificity in western blotting experiments. Endogenous SUV39H1 (45 kDa) is undetectable in mESCs, probably due to low expression levels. SUV39H1 antibody detects specific bands in western blotting experiments using U2OS cells transiently transfected with GFP-tagged Suv39h1 (72 kDa) or HA-tagged SUV39H1 (46 kDa). This confirms that the SUV39H1 antibody used in this study specifically detects SUV39H1 in the context of SUV39H1 enrichment such as SUV39H1 overexpression or immunoprecipitation (see Fig. 4a, b). **(B)** *Oct4* promoter ChIP analysis using H3K9me3 specific antibodies was performed in *Oct4P4* overexpressing mESCs. *Oct4P4* overexpressing mESCs were transiently transfected with indicated siRNAs. Knock-down of SUV39H1 decreases abundance of H3K9me3 at the promoter of the ancestral *Oct4* gene. n, refers to the number of independent experiments carried out; error bars indicate standard deviation; a Student's t-test was used to calculate statistical significance values: *, $p < .05$.

Supplementary Figure 6.



(A) Quantitative RT-PCR determining *Oct4P4* expression levels in pMEFs transiently transfected with two different siRNAs (#1, #2) targeting *Oct4P4*. Control cells were transfected with two different siRNA control oligos. *Oct4P4* expression was normalized to *gapdh*. Both *Oct4P4* specific RNAi oligonucleotides (#1, #2) significantly decrease *Oct4P4* expression levels. **(B)** Quantitative RT-PCR determining *Oct4* expression levels in pMEFs transiently transfected with two different siRNAs (#1, #2) targeting *Oct4P4*. Control cells were transfected with two different siRNA control oligos (#1, #2). *Oct4P4* expression was normalized to *gapdh*. Both *Oct4P4* specific siRNAs significantly increase ancestral *Oct4* expression. **(C)** Quantitative RT-PCR determining mESCs self-renewal marker genes expression levels in pMEFs transiently transfected with indicated siRNAs. siOct4P4#2 significantly increases ancestral Sox2, Nanog, Klf4 and hTERT expression. These results reproduce data from Fig. 2G **(D)** Measurement of telomere repeat content by quantitative RT-PCR. Total genomic DNA from pMEFs transiently transfected with indicated RNAi oligonucleotides subjected to quantitative RT-PCR using telomere specific PCR primers. Alu-equivalent B1 repeats were used as reference sequence. Transfection of siOct4P4#2 increases telomere repeat content. n, refers to the number of independent experiments carried out; error bars indicate standard deviation; a Student's t-test was used to calculate statistical significance values: *, $p < .05$; **, $p < .01$

Supplementary Figure 8.



Subcellular localization of the human OCT4 pseudogene 3 (*OCT4P3*) lncRNA. Percentage of total RNA found in the nuclear and cytoplasmic fractions of ovarian cancer cell lines OVCAR-3 and SKOV-3, as determined by quantitative real time (RT-PCR). OCT4P3 lncRNA shows prevalent nuclear localization.

Supplementary Figure 9.

▪

A

Fig. 1b Oligo random

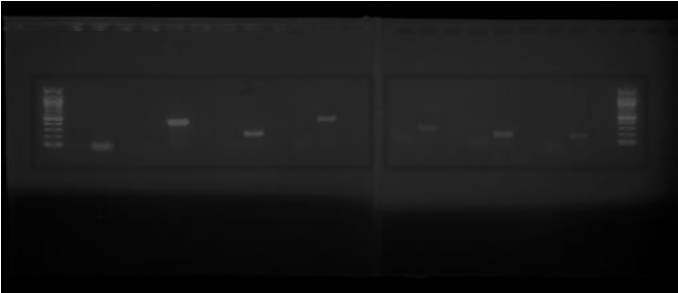


Fig. 1b Oligo dT



B

Fig. 2c OCT4



Fig. 2c ACTIN



Fig. 2f OCT4

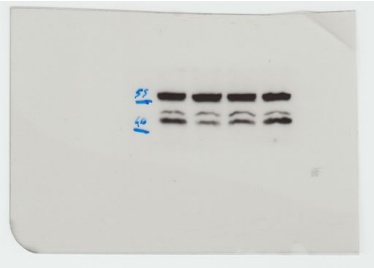


Fig. 2f ACTIN



(A) Uncropped gels from Figure 1. (B) Uncropped blots from Figure 2. Images of the selected portions shown in Figure 1 and Figure 2 are indicated.

C

Fig. 3b OCT4



Fig. 3b HA

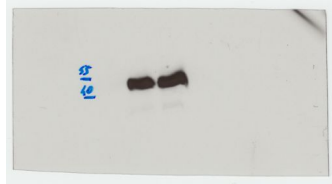


Fig. 3b ACTIN

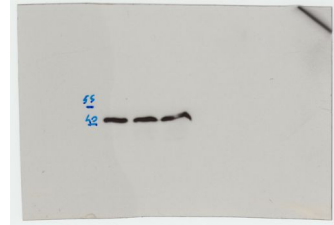


Fig. 3d OCT4

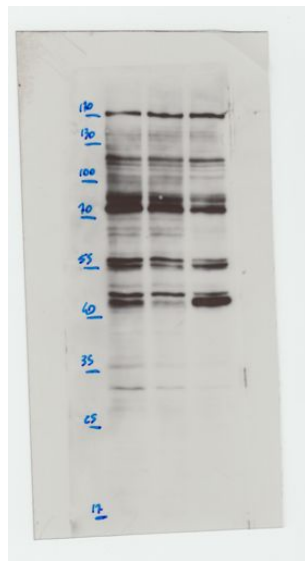


Fig. 3d ACTIN



Fig. 3e MS2-flag

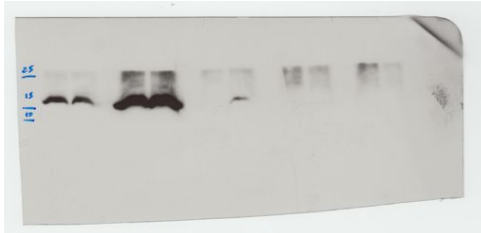
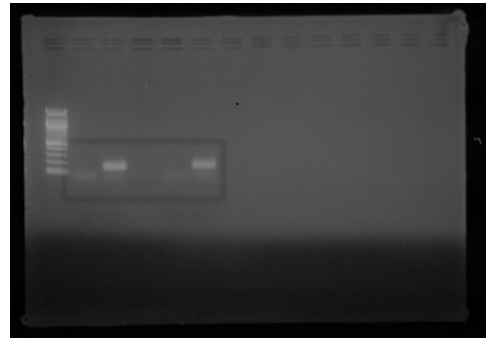


Fig. 3e Oct4P4



(C) Uncropped blots and gels from Figure 3. Images of the selected portions shown in Figure 3 are indicated.

D

Fig. 4a SUV39H1



Fig. 4a MS2-flag

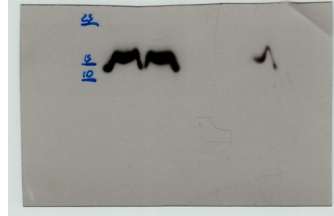


Fig. 4a Oct4P4

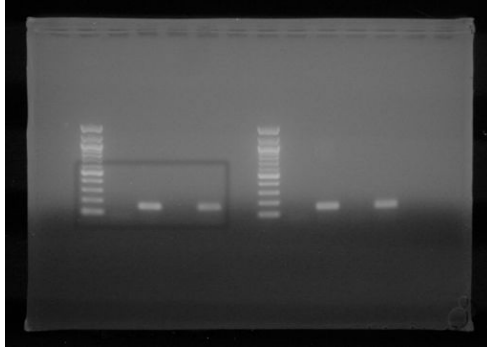


Fig. 4b SUV39H1

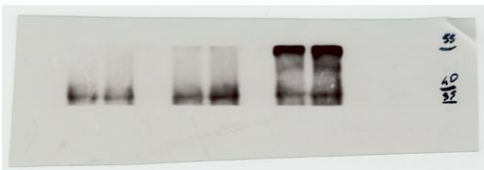


Fig. 4b Oct4P4

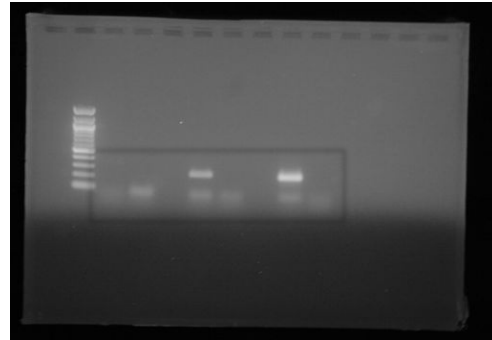


Fig. 4c OCT4

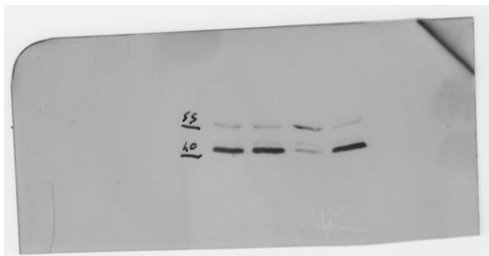
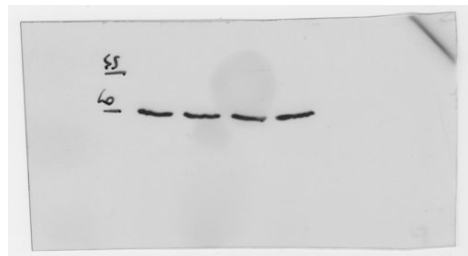


Fig. 4c ACTIN



(D) Uncropped blots and gels from Figure 4. Images of the selected portions shown in Figure 4 are indicated.

Supplementary Table 1

PCR Primers

Name	Sequence (5'-3')
Brachyury f	CAGCCCACCTACTGGCTCTA
Brachyury r	GAGCCTGGGGTGATGGTA
Fgf5 f	GTTTCCAGTGGAGCCCTT
Fgf5 r	GAGACACAGCAAATATTTCCAAAA
Gapdh f	TTCACCACCATGGAGAAGGC
Gapdh r	CCCTTTTGGCTCCAC
Gdf3 f	TCAGCTTCTCCCAGACCAGGGTTT
Gdf3 r	CACACGCCCCGGTCCTGAAC
Klf4 f	CCGGTCCCTAGAGGCCATTT
Klf4 r	AGTTCATCGGAGCGGGCGAA
Nanog f	TTCTTGCTTACAAGGGTCTGC
Nanog r	AGAGGAAGGGCGAGGAGA
Nestin f	CTGCAGGCCACTGAAAAGTT
Nestin r	GACCCTGCTTCTCCTGCTC
Oct4 f	CAGGGACACCTTTCCCAGGG
Oct4 r	TTTAAGAACAAAATGATGAG
Oct4P1 f	GGGATGCTGTGAGCCAAG
Oct4P1 r	TTTATTTAAGGATGAGTGAC
Oct4P2 f	CGTATGCTCAAGTTTCAC
Oct4P2 r	TTTAAGAACAAAATGATGAG
Oct4P3 f	TGTCACTACCCTGGCTCTC
Oct4P3 r	TGATGAGTGATGGACGAGC
Oct4P4 f	TGGCACCTGGCTTTAGACTTT
Oct4P4 r	CCAGGCCAACTTAGGGCATT
Oct4P5 f	CGGTGCCAGAAGGGCCAACG
Oct4P5 r	AGCTCCCATAGCCTGGGGTATCA
Otx f	AGCTGCTCCCTCCGAAGCAGTA
Otx r	AGCCGGACGGTCTCGATTCCG
Sox2 f	TGCTGCCTCTTTAAGACTAGGG
Sox2 r	TCGGGCTCCAACTTCTCT
Suv39h1 f	CGGATCACCGTGGAGAAT
Suv39h1 r	CACTCACAGCCAACAGCTACCT
Tert f	GCAGGTGAACAGCCTCCAGACAG
Tert r	TCCTAACACGCTGGTCAAAGGGAAGC
Oct4P4 semi-qPCR f	CTGGCACCTGGCTTTAGACTTT
Oct4P4 semi-qPCR r	CCAAGCCTTGTCCACCCTCTG
Oct4P4 primer 1 f	CTGGCACCTGGCTTTAGACTTT
Oct4P4 primer 1 r	CCAGGCCAACTTAGGGCATT
Oct4P4 primer 2 f	GTCCCTAGGTGACCAACT
Oct4P4 primer 2 r	CCAAGCCTTGTCCACCCTCTG
OctP4-5'UTR f	GTCCCTAGGTGACCAACT
OctP4-5'UTR r	GGGGAAGTTGGGCACCCC
OctP4-3'UTR f	TAGAGAGGAAGATGAAGT
OctP4-3'UTR r	TGTGTCCCAGGCTTTTTTA
OctP4-5'UTR-3'UTR f	GTCCCTAGGTGACCAACT

OctP4-5'UTR-3'UTR r TGTGTCCCAGGCTTTTTTA
hOCT4P3 f CTTCGGATTTCGCCTTCTCA
hOCT4P3 r GGCACTAGCCCCACTCCAGT

Supplementary Methods:

Plasmids

pCDNA3-Oct4P4 was generated by amplifying a genomic region encoding *Oct4P4* from mESCs by PCR (Forward: GGGAATTCCTCGAGGTCCCTAGGTGACCAACTCCT; Reverse: GGGAATTCTCTAGATGTGTCCCAGGCTTTTTTAAAT) and cloning into pCDNA3 vector (Promega) via XhoI and XbaI restriction sites. pcDNA3-Oct4p4Δ1 was generated by amplifying the corresponding region by PCR (Fragment 1: Forward: GGGAATTCCTCGAGGTCCCTAGGTGACCAACTCCT ; Reverse: GGGAATTCGAATTCCTGGCCCCATTCCAGGCCCA. Fragment 2: Forward: GGGAATTCGAATTCCTGGAGACTTTGCAGCCTTA; Reverse: GGGAATTCTCTAGATGTGTCCCAGGCTTTTTTAAAT) and insertion into XhoI and XbaI sites of pCDNA3 (Promega). pcDNA3-Oct4P4Δ2 was obtained by amplifying the relevant genomic region by PCR (Forward: GGGAATTCCTCGAGATGGCCTGGCACCTGGCTTTAG; Reverse: GGGAATTCTCTAGACCCTTTCTGTGCTTTGAACCT) and cloning into XhoI and XbaI sites of pcDNA3 (Promega). pcDNA3-Oct4P4Δ5'UTR and pcDNA3-Oct4P4Δ3'UTR were generated by amplifying the corresponding regions by PCR (pcDNA3-Oct4P4Δ5'UTR: Forward: GGGAATTCCTCGAGATGGCCTGGCACCTGGCTTTAG; Reverse: GGGAATTCTCTAGATGTGTCCCAGGCTTTTTTAAAT) (pcDNA3-Oct4P4Δ3'UTR: Forward: GGGAATTCCTCGAGGTCCCTAGGTGACCAACTCCT; Reverse: GGGAATTCTCTAGACCCTTTCTGTGCTTTGAACCT) and cloning into pcDNA3 (Promega) via XhoI and XbaI restriction sites.

pcDNA3-Oct4P4-5'UTR was obtained by PCR amplifying the 5'UTR region of Oct4P4 (Forward: AAGCTTGTCCCTAGGTGACCAACT; Reverse: CTCGAGGGGAAGTTGGGCACCCC) and insertion into HindIII and XhoI sites of pcDNA3 vector (Promega). pcDNA3-Oct4P4-3'UTR was obtained by amplifying the 3'UTR of Oct4P4 by PCR (Forward: CTCGAGTAGAGAGGAAGATGAAGT; Reverse: TCTAGATGTGTCCCAGGCTTTTTTA) and cloning into XhoI and XbaI sites of pcDNA3 (Promega). **pcDNA3-Oct4P4-5'UTR-3'UTR** was PCR-amplified from the corresponding

region of Oct4P4 (Fragment 1: Forward: AAGCTTGTCCCTAGGTGACCAACT; Reverse: CTCGAGGGGGAAGTTGGGCACCCC. Fragment 2: Forward: CTCGAGTAGAGAGGAAGATGAAGT; Reverse: TCTAGATGTGTCCCAGGCTTTTTTA) and inserted into pcDNA3 (Promega) via HindIII and XbaI .

For Suv39h1 constructs, Suv39h1 was PCR-amplified from mESC cDNA (pLPC-HA-Suv39h1: Forward: GGATCCATGGCGGAAAATTTAAAAGGT; Reverse: CTCGAGCTAGAAGAGGTATTTTCGGC) (pLPC-GFP-Suv39h1: Forward: GAATTCATGGCGGAAAATTTAAAAGGT; Reverse: CTCGAGCTAGAAGAGGTATTTTCGGC) and cloned into pLPC-HA vector via BamHI and XhoI, and into pLPC-GFP vector via EcoRI and XhoI restriction sites.

Quantitative determination of *Oct4P4* and *Oct4* RNA molecules

To estimate the copy number of endogenous *Oct4P4* and *Oct4* transcripts, vectors containing subcloned fragments of *Oct4P4* and *Oct4* (pTOPO-*Oct4P4* and pGL3-3'UTR-*Oct4*) were used as internal standards for copy number (Supplementary Reference ¹). 0.5 µg of total RNA from mESCs and pMEFs was subjected to reverse transcription using the Quantitect reverse transcription kit (Qiagen) according to the manufacturer's protocol. Equal volumes (5µl) of target cDNA and serially dilutions of vectors (mol/µl; external standard) were used in quantitative RT-PCR in order to determine the molar range of the two target RNAs. To more precisely define the number of *Oct4* and *Oct4P4* molecules in 5µl of cDNA, limited molar dilutions of vectors (internal standard) were then added to target cDNA in each quantitative RT-PCR reaction. This enables us to determine the number of *Oct4/Oct4P4* mRNA molecules in each reaction tube. The number of molecules obtained for *Oct4P4* and *Oct4* was expressed as molecules per µg of total RNA used for RT.

Embryoid body differentiation

For mESC differentiation medium (Dulbecco's Eagle's medium (DMEM) supplemented with 15% ES cell certified serum (Invitrogen), 1% non-essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), 1% L-glutamine (Invitrogen), 0.1 mM β-mercaptoethanol and 1% penicillin/streptomycin (Invitrogen) was used. 300 mESCs were cultured in hanging drops for three days. After 3 days EBs were transferred to low-attachment 24-well plates (Euroclone) and grown in suspension for 4 days.

Alkaline phosphatase assays

mESC colonies were stained using the StemTAG Alkaline Phosphatase Staining Kit (CBA-302; Cell Biolabs) following the manufacturer's instructions. Self-renewal colonies stain positive for alkaline phosphatase (AP); differentiated colonies show lower staining intensity or are negative for AP activity. For the photospectrometric measurement of AP activity in mESCs, p-nitrophenol levels were determined using StemTAG Alkaline Phosphatase Activity Assay Kit (CBA-302; Cell Biolabs) following the manufacturer's protocol.

RNA expression analysis

For semi-quantitative PCR, total RNA from ES cells was purified using Qiazol lysis reagent (Qiagen). 1.5 µg of total RNA was treated with DNase (RQ1, Promega) and subjected to reverse transcription in the presence of random primers (Promega) or oligo dT primers using SuperScriptIII Reverse Transcriptase (Invitrogen) according to the manufacturer's suggestions. The obtained cDNA was amplified by PCR on a 2720 Thermal Cycler PCR machine (Applied Biosystems) using AccuPrime Taq DNA Polymerase (Invitrogen) according to the manufacturer's instructions. PCR products were run on a 1.2% agarose gel and visualized by ethidium bromide staining. Specific PCR primers are shown in Supplementary table 1.

To detect antisense *Oct4* and *Oct4P4* transcripts, strand-specific reverse transcription polymerase chain reaction was performed. Gene-specific sense primers were used for reverse transcription; ***Oct4* asRNA: TCGATAGATGACACACAGGG; *Oct4P4* asRNA: GTCACGAATTCTCTCCGGGA**), and **gene-specific antisense primers were used to synthesize cDNAs to the sense RNA (*Oct4* senseRNA: AGCTATCTACTGTGTGTCCC; *Oct4P4* senseRNA: CAGTGCTTAAGAGAGGCCCT)**. cDNA was synthesized using the Quantitect reverse transcription kit (Qiagen) according to the manufacturer's protocol. cDNA fragments were PCR amplified using a 2720 Thermal Cycler (Applied Biosystems) using AccuPrime Taq DNA Polymerase (Invitrogen) according to the manufacturer's instructions. PCR products were run on a 1.2% agarose gel and visualized by ethidium bromide staining. **Specific PCR primers used for the amplification of both sense and antisense *Oct4* (*Oct4* f; *Oct4* r) and *Oct4P4* (*Oct4P4* f; *Oct4P4* r) transcripts are reported in Supplementary table 1.**

For quantitative real-time PCR, total RNA was extracted from cells using Qiazol lysis reagent (Qiagen) according to manufacturer's suggestions. Nuclear and cytoplasmic RNA fractions were obtained by the following protocol: cells were collected, resuspended in

