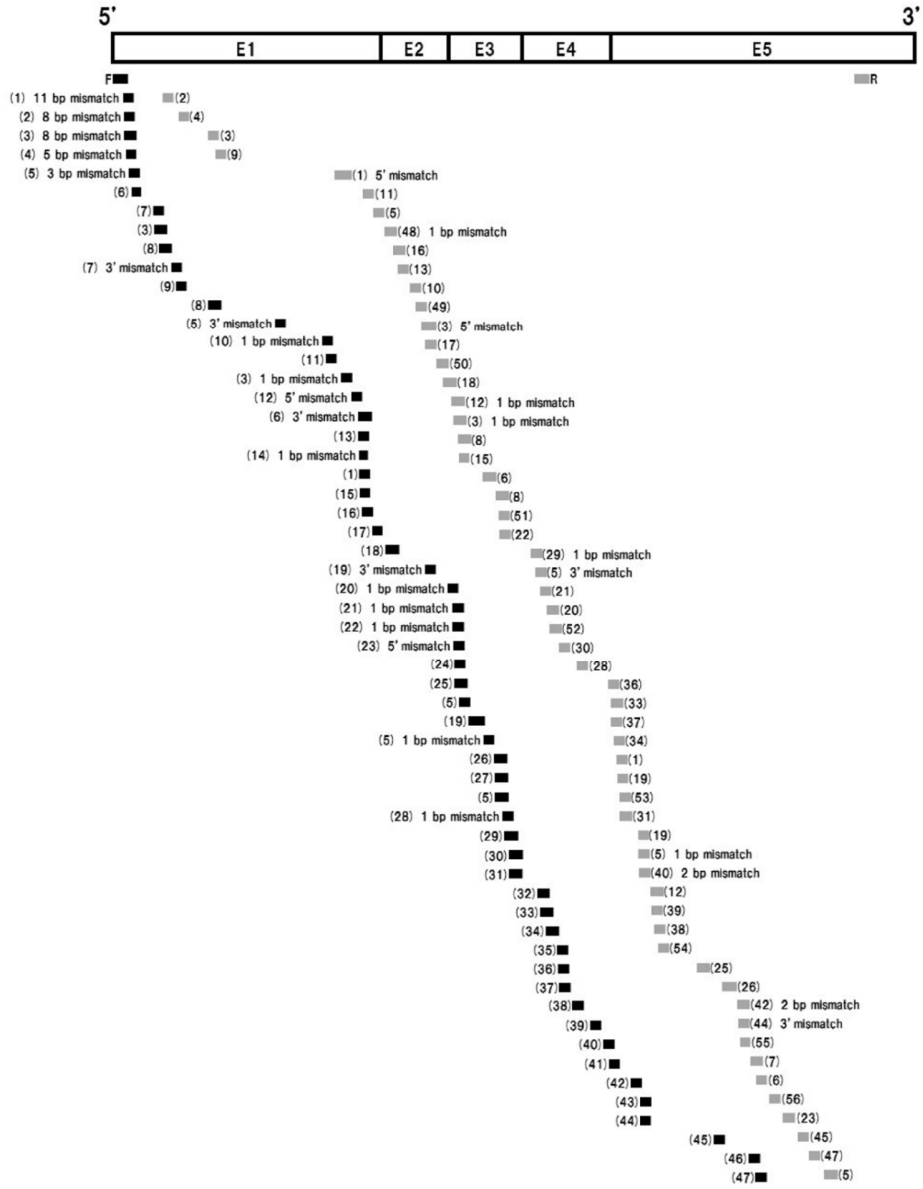


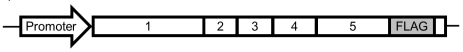
#	Refferences
1	Langenfeld E, Deen M, Zachariah E, *et al.* Mol Cancer. 2013 Oct 26; 12 (1): 129.
2	Nishimura N, Hartomo TB, Pham TV, *et al.* Environ Health Prev Med. 2012 May; 17 (3): 246-51.
3	Xu G, Yang L, Zhang W, *et al.* Stem Cells Dev. 2015 Jul 1; 24 (13): 1546-57.
4	Asadi MH, Mowla SJ, Fathi F, *et al.* Int J Cancer. 2011 Jun 1; 128 (11): 2645-52.
5	Kaltz N, Funari A, Hippauf S, *et al.* Stem Cells. 2008 Sep; 26 (9): 2419-24.
6	Mueller T, Luetzkendorf J, Nerger K, *et al.* Cell Mol Life Sci. 2009 Feb; 66 (3): 495-503.
7	Zhao S, Yuan Q, Hao H,*et al.* J Pathol. 2011 Apr; 223 (5): 672-82.
8	Gazouli M, Roubelakis MG, Theodoropoulos GE, *et al.* Mol Carcinog. 2012 Feb; 51 (2): 165-73.
9	Ball SG, Worthington JJ, Canfield AE, *et al.* Stem Cells. 2014 Mar; 32 (3): 694-705.
10	Cho YH, Han KM, Kim D, *et al.* Stem Cells. 2014 Feb; 32 (2): 424-35.
11	Yu CH, Yu CC. PLoS One. 2014 Jan 24; 9 (1): e87129.
12	Kumar M, Allison DF, Baranova NN, *et al.* PLoS One. 2013 Jul 30; 8 (7): e68597.
13	Galatro TF, Uno M, Oba-Shinjo SM, *et al.* PLoS One. 2013 Apr 16; 8 (4): e61605.
14	Mathieu J, Zhang Z, Nelson A, *et al.* Stem Cells. 2013 Sep; 31 (9): 1737-48.
15	Zhou X, Zhou YP, Huang GR, *et al.* Int J Gynecol Pathol. 2011 May; 30 (3): 262-70.
16	Saigusa S, Tanaka K, Toiyama Y, *et al.* Ann Surg Oncol. 2009 Dec; 16 (12): 3488-98.
17	Lengerke C, Fehm T, Kurth R, *et al.* BMC Cancer. 2011 Jan 28; 11: 42.
18	Clark AT, Rodriguez RT, Bodnar MS, *et al.* Stem Cells. 2004; 22 (2): 169-79.
19	Chen Z, Xu WR, Qian H, *et al.* J Surg Oncol. 2009 Jun 1; 99 (7): 414-9.
20	Zhang ZN, Chung SK, Xu Z, *et al.* Stem Cells. 2014 Jan; 32 (1): 157-65.
21	Li Z, Liu C, Xie Z, *et al.* PLoS One. 2011; 6 (6): e20526.
22	Shi J, Shi W, Ni L, *et al.* Oncol Rep. 2013 Jul; 30 (1): 201-6.
23	Leung EL, Fiscus RR, Tung JW, *et al.* PLoS One. 2010 Nov 19; 5 (11): e14062.
24	Chen WJ, Ho CC, Chang YL, *et al.* Nat Commun. 2014 Mar 25; 5: 3472.
25	Windmolders S, De Boeck A, Koninckx R, *et al.* J Mol Cell Cardiol. 2014 Jan; 66: 177-88.
26	Jung JW, Park SB, Lee SJ, *et al.* PLoS One. 2011; 6 (11): e28068.
27	Meng X, Su RJ, Baylink DJ, *et al.* Cell Res. 2013 May; 23 (5): 658-72.
28	Watanabe K, Meyer MJ, Strizzi L, *et al.* Stem Cells. 2010 Aug; 28 (8): 1303-14.

#	Refferences
29	Montserrat N, Nivet E, Sancho-Martinez I, *et al.* Cell Stem Cell. 2013 Sep 5; 13 (3): 341-50.
30	Riekstina U, Cakstina I, Parfejevs V, *et al.* Stem Cell Rev. 2009 Dec; 5 (4): 378-86.
31	Monk M, Holding C. Oncogene. 2001 Dec 6; 20 (56): 8085-91.
32	Ma L, Lai D, Liu T, *et al.* Acta Biochim Biophys Sin. 2010 Sep; 42 (9): 593-602.
33	Feng S, Duan X, Lo PK, *et al.* Integr Biol. 2013 May; 5 (5): 768-77.
34	Lu Y, Loh YH, Li H, *et al.* Cell Stem Cell. 2014 Jul 3; 15 (1): 92-101.
35	Murakami A, Takahashi F, Nurwidya F, *et al.* PLoS One. 2014 Jan 28; 9 (1): e86459.
36	Padín-Iruegas ME, Herranz-Carnero M, Aguin-Losada S, *et al.* Oncol Rep. 2013 Jun; 29 (6): 2467-72.
37	Shen WW, Zeng Z, Zhu WX, *et al.* J Mol Med. 2013 Aug; 91 (8): 989-1000.
38	Lou YR, Kanninen L, Kuisma T, *et al.* Stem Cells Dev. 2014 Feb 15; 23 (4): 380-92.
39	Bier A, Giladi N, Kronfeld N, *et al.* Oncotarget. 2013 May; 4 (5): 665-76.
40	Theunissen TW, Powell BE, Wang H, *et al.* Cell Stem Cell. 2014 Oct 2; 15 (4): 471-87.
41	Vaiphei K, Sinha SK, Kochhar R. Asian Pac J Cancer Prev. 2014; 15 (8): 3519-24.
42	Wang L, Guo H, Yang L, *et al.* Mol Cell Biochem. 2013 Jul; 379 (1-2): 7-18.
43	Yan X, Luo H, Zhou X, *et al.* Oncol Rep. 2013 Dec; 30 (6): 2733-40.
44	Xu C, Xie D, Yu SC, *et al.* Cancer Res. 2013 May 15; 73 (10): 3181-9.
45	Zhang Z, Zhu Y, Lai Y, *et al.* Int J Oncol. 2013 Oct; 43 (4): 1194-204.
46	Wang XY, Penalva LO, Yuan H, *et al.* Mol Cancer. 2010 Aug 21; 9: 221.
47	Takahashi K, Tanabe K, Ohnuki M, *et al.* Cell. 2007 Nov 30; 131 (5): 861-72.
48	Karaöz E, Okçu A, Gacar G, *et al.* J Cell Physiol. 2011 May; 226 (5): 1367-82.
49	Sakaki-Yumoto M, Liu J, Ramalho-Santos M, *et al.* J Biol Chem. 2013 Jun 21; 288 (25): 18546-60.
50	Atlasi Y, Mowla SJ, Ziaee SA, *et al.* Stem Cells. 2008 Dec; 26 (12): 3068-74.
51	Wezel F, Pearson J, Kirkwood LA, *et al.* Am J Pathol. 2013 Oct; 183 (4): 1128-36.
52	Lu Y, Zhu H, Shan H, *et al.* Cancer Lett. 2013 Oct 28; 340 (1): 113-23.
53	Tai MH, Chang CC, Kiupel M, *et al.* Carcinogenesis. 2005 Feb; 26 (2): 495-502.
54	Wen K, Fu Z, Wu X, *et al.* Cancer Lett. 2013 Jun 1; 333 (1): 56-65.
55	Ma B, Lei X, Guan Y, *et al.* Oncol Rep. 2011 Jul; 26 (1): 135-43.
56	Busch C, Bareiss PM, Sinnberg T, *et al.* J Pathol. 2009 Mar; 217 (4): 589-96.

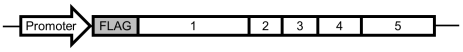
Figure S1. Schematic representation of the positions of primers used for the analysis of human *OCT4* expression in recent studies.

Black and gray boxes: forward and reverse primers, respectively. Black box F and gray box R: primers used to amplify *OCT4A* in this study. The number of base pairs indicates the difference in nucleotide number between primers and pseudogenes; the 5′ or 3′ mismatch indicates the position where a single-nucleotide mismatch exists in the 5′ or 3′ terminus between the primers and the pseudogenes. Numbers in parentheses: citation numbers corresponding to articles listed in the table.

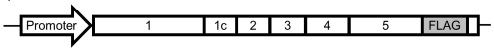
pEF1-Oct4A-FLAGc



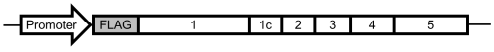
pEF1-Oct4A-FLAGn



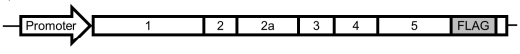
pEF1-Oct4A1-FLAGc



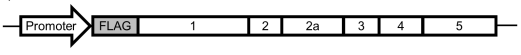
pEF1-Oct4A1-FLAGn



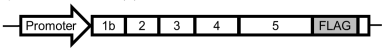
pEF1-Oct4A2-FLAGc



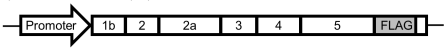
pEF1-Oct4A2-FLAGn



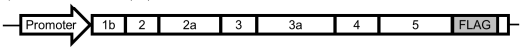
pEF1-Oct4B-FLAGc (B)



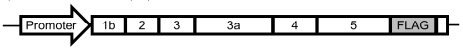
pEF1-Oct4B1-FLAGc (B1)



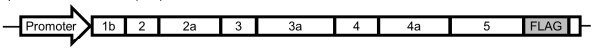
pEF1-Oct4B2-FLAGc (B2)



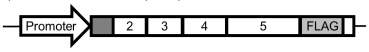
pEF1-Oct4B3-FLAGc (B3)



pEF1-Oct4Bns-FLAGc (Bns)

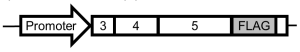


pEF1-Oct4BΔ86-FLAGc (BΔ86)

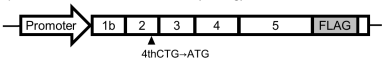


pEF1-Oct4BΔ86-E2AGG-FLAGc (BΔ86-E2agg)

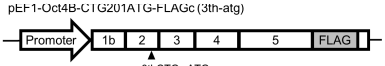
pEF1-Oct4C-FLAGc (C)



pEF1-Oct4B-CTG190ATG-FLAGc (4th-atg)

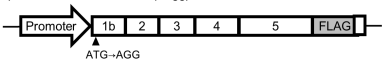


pEF1-Oct4B-CTG201ATG-FLAGc (3th-atg)

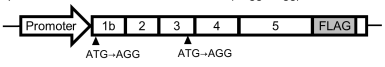


pEF1-Oct4B-ATG164AGG-FLAGc (B-Cagg)

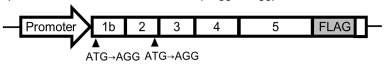
pEF1-Oct4B-AGG-FLAGc (Bagg)



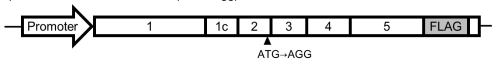
pEF1-Oct4B-AGG-ATG164AGG-FLAGc (Bagg-Cagg)



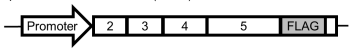
pEF1-Oct4B-AGG-E2AGG-FLAGc (Bagg-E2agg)



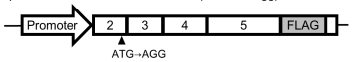
pEF1-Oct4A1-E2AGG-FLAGc (A1-E2agg)



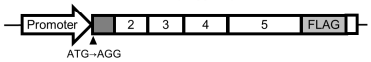
pEF1-Oct4-E2-E5-FLAGc (AΔE1)



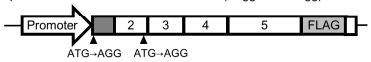
pEF1-Oct4-E2-E5-E2AGG-FLAGc (AΔE1-E2agg)



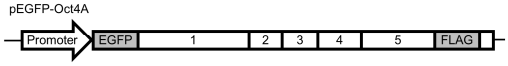
pEF1-Oct4BΔ86-AGG-FLAGc (BaggΔ86)



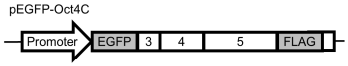
pEF1-Oct4BΔ86-AGG-E2AGG-FLAGc (BaggΔ86-E2agg)



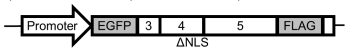
pEGFP-Oct4A



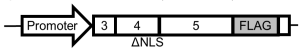
pEGFP-Oct4C



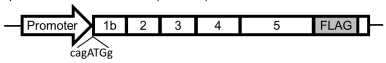
pEGFP-Oct4C-ΔNLS (EGFP-Oct4CΔNLS)



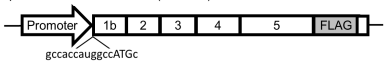
pEF1-Oct4C-FLAGc-ΔNLS (Oct4CΔNLS)



pEF1-Oct4Bnative-FLAGc (B-native)



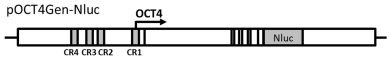
pEF1-Oct4BKozak-FLAGc (B-Kozak)



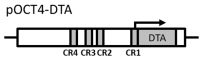
pEF1-Oct4B-E2Kozak-FLAGc (B-E2Kozak)

pOCT4Gen-FLAG

pOCT4Gen-Nluc



pOCT4-DTA



pOCT4-GFP

pPORE-Nluc

Figure S2. DNA constructs used in this study.

PCR primer set 5′-gaattcaccATGGCGGGACACCTGGCTTCAGATTTC-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from cDNAs isolated in this study, which were ligated into EcoRI and SalI site of EF1 promoter vector containing FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4A-FLAGc, pEF1-Oct4A1-FLAGc and pEF1-Oct4A2-FLAGc.

PCR primer set 5′-gaattcaccATGGCGGGACACCTGGCTTCAGATTTC-3′ and 5′-gtcgacTCAGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from cDNAs isolated in this study, which were ligated into EcoRI and SalI site of EF1 promoter vector containing FLAG-tag in the 5′ end of MCS to construct pEF1-Oct4A-FLAGn, pEF1-Oct4A1-FLAGn and pEF1-Oct4A2-FLAGn.

PCR primer set 5′-gaattcAGGCAGATGCACTTCTACAGACTATTC-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from cDNAs isolated in this study, which were ligated into the EcoRI and SalI site of the EF1 promoter vector containing FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4B-FLAGc, pEF1-Oct4B1-FLAGc, pEF1-Oct4B2-FLAGc, pEF1-Oct4B3-FLAGc and pEF1-Oct4Bns-FLAGc.

PCR primer set 5′-TCCCAGGACATCAAAGCTCTG-3′ and 5′-GTGTGGCCCCAAGGAATAGTC-3′ was utilized to construct pEF1-Oct4BΔ86-FLAGc from pEF1-Oct4B-FLAGc.

PCR primer set 5′-CCGAGGTGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-Oct4BΔ86-E2AGG-FLAGc from pEF1-Oct4BΔ86-FLAGc.

PCR primer set 5′-gaattcaccATGTGTAAGCTGCGGCCCTTGCT-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from pEF1-Oct4A-FLAGc, which were ligated into the EcoRI and SalI site of the EF1 promoter vector containing FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4C-FLAGc.

PCR primer set 5′-ATGGGGGTTCTATTTGGTGG-3′ and 5′-AGCTTTGATGTCCTGGGACT-3′ was utilized to construct pEF1-Oct4B-CTG190ATG-FLAGc from pEF1-Oct4B-FLAGc. PCR primer set 5′-ACCAGGGGATATACACAGGC-3′ and 5′ -GATCCTCTTCTGCTTCAGGA-3′ was utilized to construct pEF1-Oct4B-CTG201ATG-FLAGc from pEF1-Oct4B-FLAGc.

PCR primer set 5′-AGGTGTAAGCTGCGGCCCTTGCTGCAG-3′ and 5′-GTTCTTGAAGCTAAGCTGCAGAGCCTC-3′ was utilized to construct pEF1-Oct4B-ATG164AGG-FLAGc from pEF1-Oct4B-FLAGc.

PCR primer set 5′-gaattcAGGCAGAGGCACTTCTACAGACTATTC-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from pEF1-Oct4B-FLAGc, which were ligated into the EcoRI and SalI site of the EF1 promoter vector containing a FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4B-AGG-FLAGc.

PCR primer set 5′-AGGTGTAAGCTGCGGCCCTTGCTGCAG-3′ and 5′-GTTCTTGAAGCTAAGCTGCAGAGCCTC-3′ was utilized to construct pEF1-OCT4B-AGG-ATG164AGG-FLAGc from pEF1-Oct4B-ATG164AGG-FLAGc.

PCR primer set 5′-CCGAGGTGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-OCT4B-AGG-E2AGG-FLAGc from pEF1-Oct4B-AGG-FLAGc.

PCR primer set 5′-CCGAGGTGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-Oct4A1-E2AGG-FLAGc from pEF1-Oct4A1-FLAGc.

PCR primer set 5′-gaattcTCCCAGGACATCAAAGCTCTG-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAG-3′ was utilized to amplify PCR fragments from pEF1-Oct4A-FLAGc, which were ligated into the EcoRI and SalI sites of the EF1 promoter vector containing a FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4-E2-E5-FLAGc.

PCR primer set 5′-CCGAGGTGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-Oct4-E2-E5-E2AGG-FLAGc from pEF1-Oct4-E2-E5-FLAGc.

PCR primer set 5′-gaattcAGGCAGAGGCACTTCTACAGACTATTC-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from pEF1-

Oct4BΔ86-FLAGc, which were ligated into the EcoRI and SalI sites of the EF1 promoter vector containing a FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4BΔ86-AGG-FLAGc.

PCR primer set 5′-CCGAGGTGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-Oct4BΔ86-AGG-E2AGG-FLAGc from pEF1-Oct4BΔ86-AGG-FLAGc.

PCR primer set 5′-gaattcaccATGGTGAGCAAGGGCGAGGA-3′ and 5′-gaattcTTTCTTGTACAGCTCGTCCATGC-3′ was utilized to amplify the ORF of EGFP, which was ligated into the EcoRI site of pEF1-Oct4A-FLAGc and pEF1-Oct4C-FLAGc to construct pEGFP-Oct4A and pEGFP-Oct4C, respectively.

PCR primer set 5′-GCTGCCGCTACCAGTATCGAGAACCGAGTGA-3′ and 5′-TGCAGCGGCCTGCACGAGGGTTTCTGCTTTG-3′ was utilized to construct pEGFP-Oct4C-ΔNLS and pEF1-Oct4C-FLAGc-ΔNLS from peg-OCTC and pEF1-Oct4C-FLAGc respectively.

PCR primer set 5′-gaattcGCCACCATGGCCATGCACTTCTACAGACTATTC-3′ and 5′-gtcgacCGTTTGAATGCATGGGAGAGC-3′ was utilized to amplify PCR fragments from pEF1-Oct4B-FLAGc, which were ligated into the EcoRI and SalI sites of the EF1 promoter vector containing a FLAG-tag in the 3′ end of the MCS to construct pEF1-Oct4BKozak-FLAGc.

PCR primer set 5′-ACCATGGGGGGCTCACCCTGGGGGTTCTA-3′ and 5′-CCTGTGTATATCCCAGGGTGATCCTC-3′ was utilized to construct pEF1-Oct4B-E2Kozak-FLAGc from pEF1-Oct4B-FLAGc.

Pou5f1 genomic DNA fragment containing the -5000 bp to +10784 bp region was subcloned from BAC clone RPCI-11-1058J10 by using a BAC subcloning kit (GENE BRIDGE) and the following primers: 5′-CTAAACTTGGCCTCTAGTGGCCTCAGAACATCCCCACTGTGATATCCGTCGACTCTAGAGGGCCCA-3′ and 5′-GTGCCCTCAGCTCCTCCACCTTCCCCACATGAACCAGTCCGCACCTATCAGAAATCGTCGACAAGCTTGATGCAT-3′. Cloned construct was designated as pOCT4Gen. KpnI fragment of pOCT4Gen was exchanged with a KpnI fragment that contained FLAG or Nluc (Promega) tagged to the 3′ end of exon 5 ORF designated as pOCT4Gen-FLAG and pOCT4Gen-Nluc. pOCT4Gen-FLAG vector backbone contained CMV RFP to visualize transfected cells. NcoI fragment of pOCT4Gen (containing -4733bp to +77bp) was ligated into the NcoI site of diphtheria toxin fragment A (DTA) vector to construct pOCT4-DTA and EGFP vector to construct pOCT4-GFP respectively.

6X PORE (5′-ATTTGAAATGCAAAT-3′)-Nluc was constructed by ligating Nluc ORF to the 6X PORE vector, designated as pPORE-Nluc.

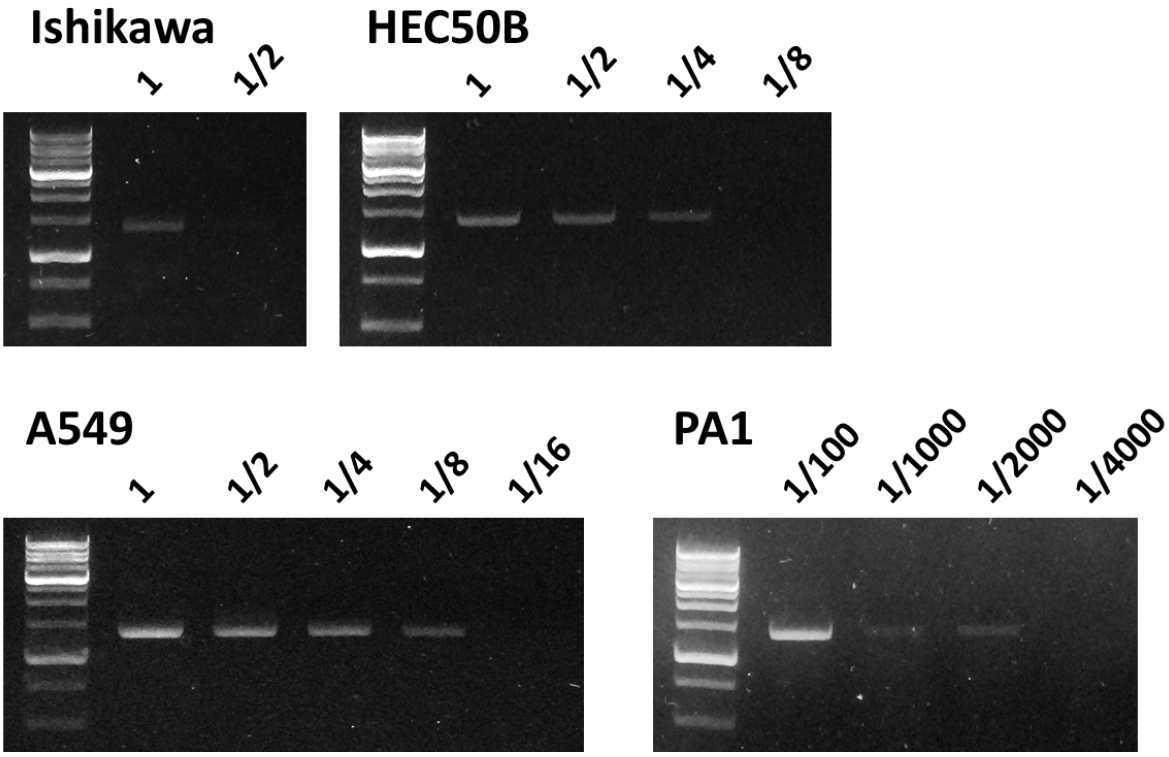


Figure S3. *OCT4A* expression levels by semi-quantitative RT-PCR.

PA-1 cDNA dilution series was established for standard PCR samples, and a 2-fold cDNA dilution series for cancer cell line samples was established for comparison of PCR samples. Typical data are shown. DNA marker: 1-kb ladder marker.

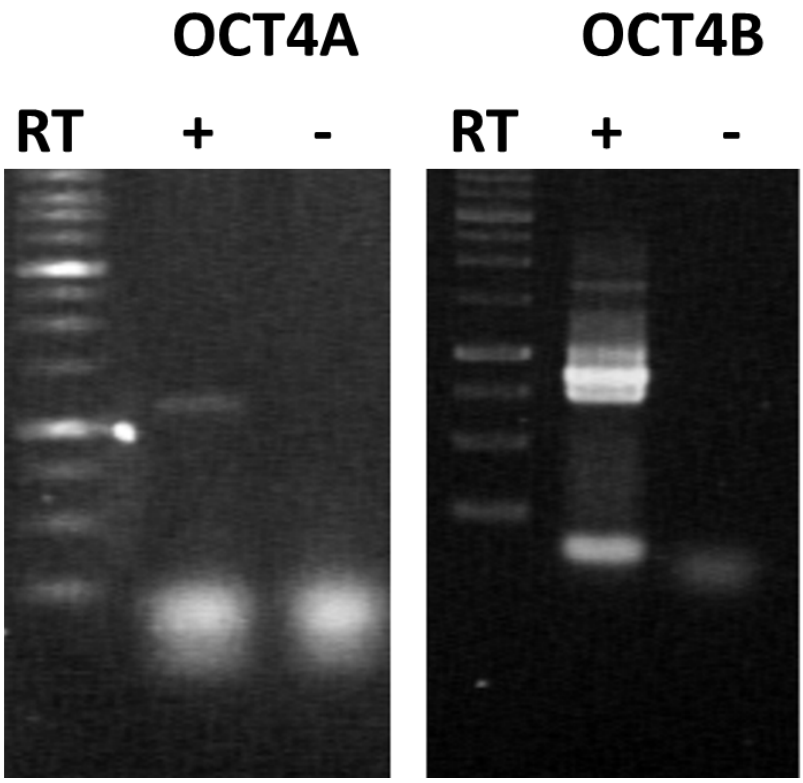


Figure S4. Detection of transcripts from pOCT4Gen-FLAG construct.

Total mRNA extracted by using Trizol reagent from the pOCT4Gen-FLAG construct introduced into A549 cells. Reverse transcription was performed by using the oligo-dT primer 5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT-3' and superscript III reverse transcriptase. *OCT4A* expression was detected by using the 5'-CTACTTGTCGTCATCGTCCTTGTAATC-3' primer and set A1 forward primer. Nested PCR was performed to detect *OCT4B* expression. The first PCR was performed by using 5'-GACTCGAGTCGACATCGA-3' primer and set B forward primer. Second PCR is performed by using 5'-CTACTTGTCGTCATCGTCCTTGTAATC-3' primer and set B forward primer.

RT+ and RT- indicate the presence or absence of reverse transcriptase treatment, respectively. DNA marker: 1-kb ladder marker.

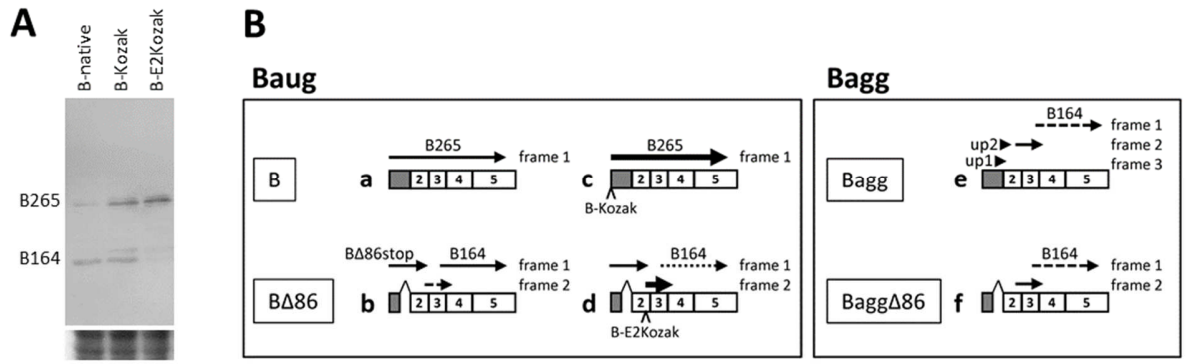


Figure S5. Schematic model of translation products from the *OCT4B* construct.

A) Western blot analysis of the *OCT4B* construct expressed in COS7 cells. B-Kozak indicates the AUG context in exon 1b with a strong Kozak sequence [equivalent to (Bc)]. B-E2Kozak indicates the AUG context of the out-of-frame AUG codon present in exon 2 with a strong Kozak consensus sequence.

B) Schematic diagram of the transcripts and translation products. Bold Baug and Bagg indicate DNA constructs harboring in-frame AUG or AGG codons in exon 1b, respectively. Two transcripts were generated from the Baug construct (Baug and BaugΔ86; a and b) and Bagg construct (Bagg and BaggΔ86; e and f). The thickness of the arrows indicates the amount of translation products. A broken arrow indicates less production as compared with the solid arrow, and a dotted arrow indicates less production than the broken arrow. Frames 1, 2, and 3 indicate three ORFs.

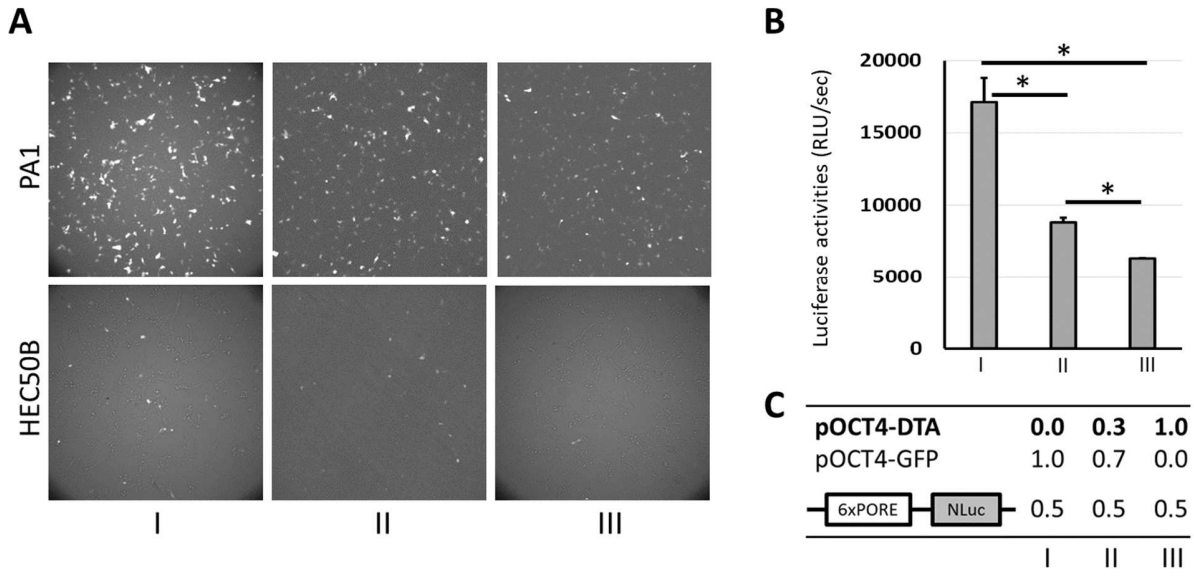


Figure S6. Bioluminescence imaging and luciferase assay using the pPORE-Nluc reporter.

PA1 and HEC50B cells were transfected with pPORE-Nluc with pOCT4-DTA and/or pOCT4-GFP. The ratio of the DNA amount of each construct per well is indicated in C, I, II, III. Positive cell numbers from bioluminescence imaging (A) and relative luciferase activity (B) were significantly decreased in a pOCT4-DTA-concentration-dependent manner. The quantitative data of luciferase activities are presented as the mean ± S.D. *; $p < 0.001$, Student's *t*-test.

Notably, it is important that the most suitable amount of pPORE-Nluc reporter DNA be used for each cell line. In HEC50B cells, $5x10^4$ cells per well were inoculated into 24-well plates and 0.5 μg of pPORE-Nluc reporter DNA per well was used for transfection in this experiment.

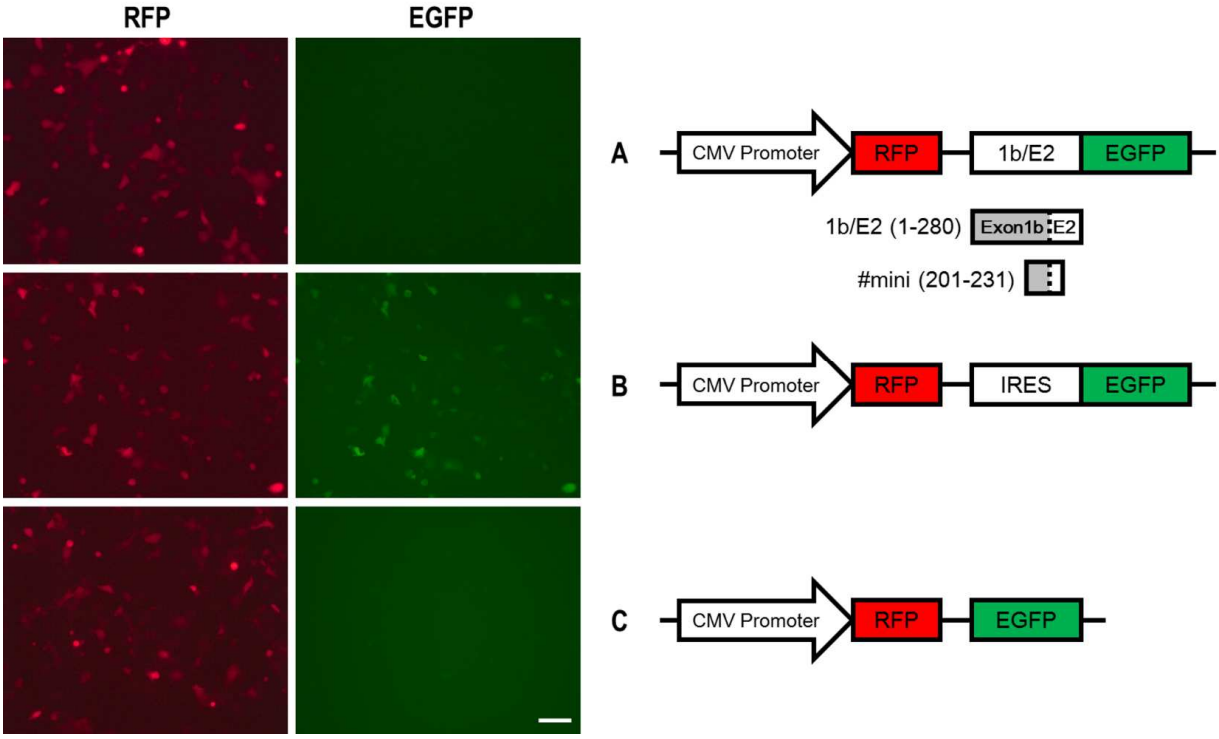


Figure S7. Negation of the alternative translation of OCT4B by an IRES.

COS7 cells were transiently transfected with bicistronic plasmids, and RFP and EGFP fluorescence were observed (left panel). Scale bar; 50 μm. The plasmids are schematically shown (right panel).

A) We constructed and tested the 1b/E2 region as a putative IRES element (28) containing the minimal IRES elements (#mini (201–231)) reported previously (51). A dotted line represents the exon 1b/exon 2 boundary located between nt 219 and 220. EGFP signals were not detected.

B) Positive control, CMV-RFP-IRES-EGFP; EGFP signals were detected.

C) Negative control, CMV-RFP-EGFP; EGFP signals were not detected.

A RFP DNA fragment was amplified by the PCR primer set 5'-ggatccACCATGGCCTCCTCCGAGGA-3' and 5'-agatcTTAGGCGCCGGTGGAGTGGC-3', and the PCR product was ligated into the BglII site of

the pEGFP-N1 vector to construct the CMV-RFP-MCS-EGFP vector and into the pIRES2-EGFP vector to construct the CMV-RFP-IRES-EGFP vector. A partial sequence of human OCT4 containing the predicted IRES sequence was amplified using the PCR primer set 5'-gaattcTAGTCCTTTGTTACATGCATGAGTCAGTGA-3' and 5'-ggatccCTGCTTCAGGAGCTTGGCAAATTGCTC-3' (292 bp) and ligated between the EcoRI and BamHI sites of the CMV-RFP-MCS-EGFP vector, which was designated as the CMV-RFP-I1/E2-EGFP vector.

Table S1. Cells used in this study.

Cell Line	Cell Type	ATG/AGG*
Cancer cells		
MCF7	Human breast adenocarcinoma cell line	ATG
HeLa	Human endocervical adenocarcinoma cell line	ATG
Ishikawa	Human endometrial adenocarcinoma (well-differentiated) cell line	ATG/AGG
HEC265	Human endometrial adenocarcinoma (well-differentiated) cell line	ATG
HEC1	Human endometrial adenocarcinoma (moderatory differentiated) cell line	ATG
HEC50B	Human endometrial adenocarcinoma (poorly differentiated) cell line	ATG
TTA1	Human thyroid papillary carcinoma cell line	AGG
A549	Human lung adenocarcinoma cell line	ATG
S2	Human lung small cell carcinoma cell line	AGG
PA1	Human ovary teratocarcinoma cell line	ATG
Transformed cells		
HEK293T	Human embryonic kidney cell line (SV40T)	AGG
Normal cells		
ARPE-19	Human retinal pigmented epithelial cell line	ATG
HFF	Human foreskin fibroblast cell line	ATG
HUVEC	Human umbilical vein endothelial cells	ATG
HAoSMC	Human aortic smooth muscle cells	AGG

* indicates whether first-ATG of *OCT4B* is converted to AGG.

DNA polymorphism analysis

Genomic DNA was isolated from cells using standard procedures. The following PCR primer set was used to isolate a DNA fragment from genomic DNA: Forward primer 5′-CATGCATGAGTCAGTGAACAGGGAATG-3′ and reverse primer 5′-TGTAAGAACATAAACACACCAGTTATC-3′ (PCR product: 183 bp). PCR products were ligated into the pCR-Blunt vector and sequenced.

Table S2. Sequencing results for the PCR products.

Cell line	OCT4A clone #				OCT4B clone #					
	A	A1	A2	total	B	B1	B2	B3	Bns	total
MCF7	-	-	-	-	0	0	0	2	11	13
HeLa	14	1	0	15	5	1	2	0	2	10
Ishikawa	11	0	0	11	0	0	0	0	10	10
HEC265	17	6	0	23	3	8	0	3	5	19
HEC1	16	3	3	22	6	11	1	0	1	19
HEC50B	15	10	0	25	8	3	1	1	4	17
TTA1	-	-	-	-	0	0	0	0	5	5
A549	11	0	1	12	7	3	0	0	3	15
S2	-	-	-	-	0	0	0	0	15	15
PA1	22	0	0	22	4	4	1	2	9	20
HEK293T	16	0	0	16	0	0	0	0	12	12
ARPE-19	-	-	-	-	0	0	0	0	6	6
HFF	-	-	-	-	1	0	0	0	6	7
HUVEC	-	-	-	-	0	0	0	0	6	6
HAoSMC	-	-	-	-	-	-	-	-	-	-

OCT4, octamer-binding transcription factor 4.