

Expression and Differentiation between OCT4A and Its Pseudogenes in Human ESCs and Differentiated Adult Somatic Cells

Mojca Jez¹, Sakthikumar Ambady², Olga Kashpur³, Alexandra Grella³, Christopher Malcuit^{4,5✉}, Lucy Vilner⁵, Primoz Rozman¹, Tanja Dominko^{3,4,5*}

1 Blood Transfusion Centre of Slovenia, Ljubljana, Slovenia, **2** Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts, United States of America, **3** Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, Massachusetts, United States of America, **4** Bioengineering Institute, Worcester Polytechnic Institute, Worcester, Massachusetts, United States of America, **5** CellThera, Inc., Worcester, Massachusetts, United States of America

Abstract

The POU5F1 gene codes for the OCT4 transcription factor, which is one of the key regulators of pluripotency. Its transcription, alternative splicing, and alternative translation leading to the synthesis of the active, nuclear localized OCT4A has been described in detail. Much less, however, is known about actively transcribed OCT4 pseudogenes, several of which display high homology to OCT4A and can be expressed and translated into proteins. Using RT-PCR followed by pseudogene-specific restriction digestion, cloning, and sequencing we discriminate between OCT4A and transcripts for pseudogenes 1, 3 and 4. We show that expression of OCT4 and its pseudogenes follows a developmentally-regulated pattern in differentiating hESCs, indicating a tight regulatory relationship between them. We further demonstrate that differentiated human cells from a variety of tissues express exclusively pseudogenes. Expression of OCT4A can, however be triggered in adult differentiated cells by oxygen and FGF2-dependent mechanisms.

Citation: Jez M, Ambady S, Kashpur O, Grella A, Malcuit C, et al. (2014) Expression and Differentiation between OCT4A and Its Pseudogenes in Human ESCs and Differentiated Adult Somatic Cells. PLoS ONE 9(2): e89546. doi:10.1371/journal.pone.0089546

Editor: Austin John Cooney, Baylor College of Medicine, United States of America

Received September 18, 2013; **Accepted** January 22, 2014; **Published** February 24, 2014

Copyright: © 2014 Jez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was funded by the National Institutes of Health (NIH) R01GM085456 to TD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Co-authors Christopher Malcuit, Lucy Vilner and Tanja Dominko are or were employed by CellThera during the course of this study. The research presented was funded by Federal Government sources (namely NIH and DoD). Under a Memorandum of Understanding between Worcester Polytechnic Institute (WPI) and CellThera, any intellectual property arising from the work would be associated with WPI, a non-profit educational organization. CellThera does not have any commercial interest in the results of the present study. Smooth muscle cells were donated by Marsha Rolle (Lonza) and heart RNA from Glenn Gaudette (Clontech). This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: tdominko@wpi.edu

✉ Current address: Department of Biology, Kent University, Kent, Ohio, United States of America

Introduction

Octamer binding protein OCT4 (POU5F1), a POU family transcription factor transcribed from *POU5F1* gene is critical for maintenance of totipotency in primate blastomeres [1,2] and for pluripotency of the inner cell mass of developing mammalian embryos [1,3–6]. This critical role of OCT4 is maintained in embryonic stem cells [7–10], the developing germline [11–13], and germline tumors, such as seminomas and embryonic teratocarcinomas [14–18]. When inducing pluripotency in differentiated cells, OCT4 appears to be an essential, but not the only required transcription factor [19,20]. OCT4 acts as a transcriptional activator of genes involved in maintenance of undifferentiated state and as a repressor of differentiation-specific genes [7,21,22].

In addition to its presence in embryonic and germ-derived cells, expression of OCT4 and other pluripotency markers has been reported in cancers, adult stem cells and various adult somatic tissues. The most abundant reports about OCT4 expression can be found for hematopoietic stem cells and mesenchymal stem cells from umbilical cord blood and bone marrow [23–28]. Detection of OCT4 transcripts in these compartments has led to a suggestion

that embryonic OCT4A may participate in maintenance of stemness in adult stem cell compartments [29], somatic tissue cancers (bladder [30], squamous cell carcinoma [31], breast carcinoma [32,33] and may even be expressed in normal differentiated tissues [34,35]. Most astoundingly, existence of true pluripotent embryonic stem-like cells in adult organs has been proposed [23–25,36,37]. We have described expression of OCT4 in adult human fibroblasts undergoing induction into a regeneration competent cell type (iRC) using low oxygen culture and medium supplementation with FGF2 [35,38,39]. iRC cells display increased life span, maintain normal karyotype, express a number of stem cell genes, participate in *in vivo* regeneration of skeletal muscle, but do not form tumors [39]. While we observed that the amount of OCT4 transcript does not change before and after induction, we at the time did not examine whether the transcript belonged to the OCT4A or its pseudogenes.

The *POU5F1* gene produces three transcripts by alternative splicing [40,41] and three protein isoforms by alternative codon use [42] (Figure 1). Only the OCT4A transcript contains the first exon at its 5' sequence and only protein translated from this transcript consequently contains the unique N-terminal transactivation domain that regulates DNA binding of OCT4A to target

promoters. It is only this full length OCT4A - a 360 amino acid protein - that localizes to the nucleus, binds to OCT4-regulated promoters, and contributes to the maintenance of pluripotency and self-renewal [43]. OCT4A transcript and protein can be distinguished from the non-pluripotency variants (OCT4B) using primers specific for 5' sequence of OCT4A mRNA and N-terminal-recognizing antibodies, respectively, as well as by their distinct subcellular localization [41,42,44,45].

With a more recent description of OCT4 pseudogenes and their expression [46–48], however, studies examining OCT4 expression have to consider an additional level of complexity (Table S1, Figure S1). Eight OCT4 pseudogenes have been identified using bioinformatics approach [47] and several of these are expressed and translated into proteins in somatic tumors, but not in embryonic carcinomas, fibroblasts or normal tissues [48–51]. Among the OCT4 pseudogenes, OCT4-pg1, OCT4-pg3 and OCT4-pg4 pseudogenes share a high sequence homology to OCT4A and share the unique N-terminal coding sequence. Similarly, cell types isolated from umbilical cord or adult bone marrow express OCT4-pg3, OCT4-pg4, and OCT4-pg5 but do not express embryonic OCT4A [52], and reports identifying these transcripts as OCT4A may have been erroneous. The only adult, non-cancer cell type that has been reported to express pseudogenes are cells derived from peripheral blood [36]. It has been suggested that pseudogenes may play regulatory roles in expression of genes from which they originate, potentially by an RNA interference-like mechanisms [53,54] or even as weak transcriptional activators [48].

When studying expression of OCT4 it is therefore critically important to account for OCT4 transcripts, OCT4 pseudogene transcripts and all the possible protein products, the majority of which have no demonstrated DNA-binding activity. While

methods and reagents to distinguish between OCT4A and OCT4B transcripts and protein variants are available, unavailability of similar reagents to distinguish between embryonic OCT4A and its pseudogene transcripts and proteins remains an obstacle to studying activity of these genes in more detail. Here we describe unique endonuclease restriction sites within OCT4A and its pseudogenes that enable differentiation between their transcripts, and investigate their expression in undifferentiated and differentiated human cells.

We demonstrate that embryonic stem cells express exclusively OCT4A and that during differentiation this expression ceases and is replaced by expression of pseudogenes. Similar regulation, however faster, takes place in teratocarcinoma cells. We also show that all differentiated cell types express exclusively pseudogenes, but this expression can shift to predominantly embryonic OCT4 in adult human fibroblasts undergoing induction of regenerative competence (iRC) by FGF2 and low oxygen.

Materials and Methods

Cell Culture

Human embryonic stem cells (hESCs, WA09) and induced pluripotent stem cells (iPSCs, DF19-9-7T) were obtained, cultured, and maintained as recommended by the supplier (WiCell). For differentiation into embryonic bodies (EBs), stem cells were either trypsinized (hESCs) or mechanically dissociated (iPSCs) and plated onto low attachment culture plates in differentiation medium consisting of DMEM-F12, 7.5% serum replacement, 10% FBS, 0.5× NEAA (Gibco), 0.005% β-mercaptoethanol (Gibco), and 2 μmol of retinoic acid (Sigma). Medium was changed with fresh medium twice a week. On day 5 of differentiation, EBs were transferred to cell culture plates where

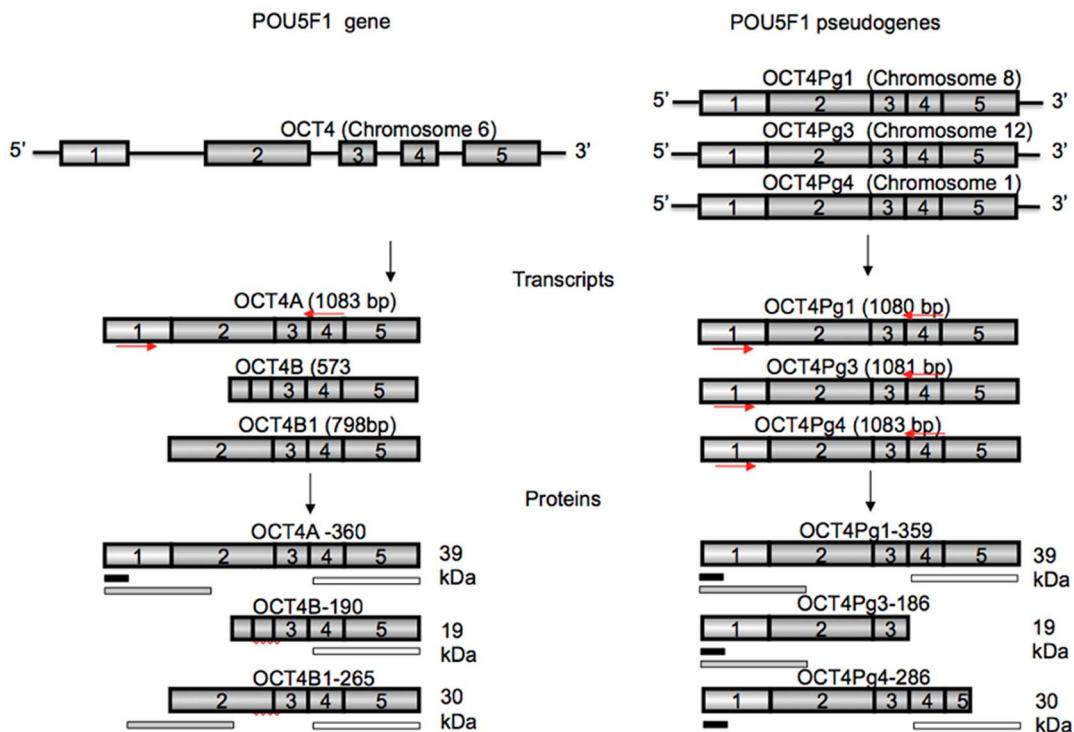


Figure 1. Schematic representation of OCT4 and its pseudogenes, transcript variants, and proteins. Red arrows indicate sequence against which the OCT4 primers were designed. Black rectangle – antigen peptide for sc5279 antibody; Grey rectangle - antigen peptide for ab19857 antibody; White rectangle - antigen peptide for ab18976 antibody.
doi:10.1371/journal.pone.0089546.g001

they were cultured until day 21. NTERA, NCCIT (both teratocarcinoma), HeLa, and NHSY (glioblastoma) cells were obtained from the American Tissue Culture Collection (ATCC). For differentiation, NTERA cells were seeded at density 1×10^6 cells per 75 cm^2 in expansion medium (DMEM +10% FBS). After attachment the medium was changed to differentiation medium consisting of DMEM +10% FBS supplemented with 0.01 mM trans-retinoic acid (Sigma). Medium was changed twice a week and samples for RNA isolation taken during 21 days post-differentiation.

Other cell types or their total RNA was obtained from Lonza (human mesenchymal stem cells – hMSCs, human umbilical cord vascular endothelium - HUVEC), Clontech (testis RNA), Lonza (smooth muscle cells, and Clontech (RNA from heart). All cell types were cultures as recommended by suppliers. Control adult primary human fibroblasts CRL-2352 (ATCC) were cultured in DMEM-F12, 10% FBS at ambient oxygen, or in 2% oxygen in the same medium, supplemented with 4 ng/mL of FGF2 (Pepro-Tech).

Total RNA Extraction and RT-PCR

Total RNA was extracted using TRIZOL reagent (Invitrogen). RNA integrity was examined on 1% agarose gel. 1 μg of total RNA was reverse transcribed using QuantiTech Reverse Transcription kit (Qiagen) according to the manufacturer's protocol. Total RNA was first treated with DNA Wipeout buffer to remove potential contamination of genomic DNA and then used for reverse transcription. The cDNA samples were stored at -20°C . The polymerase chain reaction contained 2 \times GoTaq Green Mastermix (Promega), 20 pmols of primers and PCR water. The PCR primers used for PCR were: NANOG F: 5'-TGTCTTCTGCTGAGATGCCCTCAC-3' and R: 5'-CCTTCTGCGTCACACCATTGCTAT-3'; CRIPTO1 F: 5'-GACGAGCAAATTCCCTGATGG-3' and R: 5'-GCCTCTTTCCCCCTAATTG-3'; GAPDH F: 5'-ATCAC-CATCTTCCAGGAGCGA-3' and R: 5'-TTCTCCATGGTGGTGAAGACG-3'; OCT4 F, 5'-GTTGATCCTCGGACCTGGCTA-3'; R, 5'-GGTGCCTCTCACTCGGTTCT-3'. The OCT4 primers were designed to bind at nucleotides 142–162 and 768–788 of the coding OCT4 sequence and amplify OCT4 as well as pseudogene 1, 3 and 4 transcripts (Figure 1). PCR cycling parameters consisted of an initial cDNA denaturation of 2 min at 95°C , followed by 15 sec at 94°C , 30 sec at annealing temperature 60°C , and DNA extension for 1 min at 72°C for 31 cycles. A final extension step was performed at 72°C for 10 min and held at 4°C until analysis. All PCR reactions included no template controls or no RT reaction. PCR products were resolved by 1% agarose gel electrophoresis and were visualized by exposure to 535 nm light on Kodak MM4000 image analyzer.

PCR Clean-up

Four 25 microliter reactions were combined and PCR product cleaned using NucleoSpin Extract II kit (Macherey-Nagel). Briefly, PCR reactions were diluted with 2 volumes of Buffer NT and loaded on NucleoSpin Extract II column and centrifuged 1 min at 11,000 $\times g$. Flow-through was discarded and column washed with 700 μL of buffer NT3 and then the membrane was dried for 2 min at 11,000 $\times g$. DNA was eluted into a new 1.5 mL tube and DNA concentration was determined with Nanodrop.

Restriction Digest

Sequence alignments between OCT4A and OCT4-pg1, OCT4-pg3 and OCT4-pg4 were used to identify restriction sites

that were then used for differentiation between the transcripts (Figure 2). A previously described ApaI restriction site specific for embryonic OCT4A [55] and three novel restriction sites for pseudogenes were selected (HinfI for OCT4-pg1; BglII for OCT4-pg3, XhoI for OCT4-pg4; all from NEBiolabs). 500 ng of PCR amplicons were digested with 12.5 U ApaI, 10 U XhoI, 10 U BglII, or 10 U HinfI. ApaI digestion reaction was incubated for 1.5 h at room temperature and the other three for 1.5 h at 37°C . Digested products were separated on 2% agarose gel with ethidium bromide and visualized by exposure to 535 nm light on a Kodak MM4000 image analyzer.

Cloning and Sequencing

PCR amplicons were cloned into pCR4-TOPO vector (TOPO TA Cloning Kit for Sequencing, Invitrogen), were transformed into TOPO10 competent *E. coli*, and were then plated onto LB-agar plates containing 100 $\mu\text{g}/\text{ml}$ ampicillin. Antibiotic resistant colonies were selected, expanded, and plasmid was isolated, and bi-directionally sequenced from both T3 and T7 promoters (Genewiz). Three clones were sequenced for OCT4A from ESCs, 3 clones for each pseudogene from control fibroblasts and 6 clones for OCT4A from treated fibroblasts.

Western Blotting

Total cell lysates were prepared using common procedures. Nuclear and cytoplasmic fractions were prepared as follows: cells were collected using 0.05% trypsin (Cellgro) and washed twice with PBS and pelleted. Buffer A (50 mM NaCl, 10 mM HEPES pH8.0, 500 mM sucrose, 1 mM EDTA, 0.5 mM spermidine, 0.15 mM spermine, 0.2% Triton X-100, 7 mM 2-mercaptoethanol) supplemented with complete protease inhibitor cocktail (Roche) was used to isolate cytoplasmic fractions. The cell pellets were then washed with buffer B (50 mM NaCl, 10 mM HEPES pH8.0, 25% glycerol, 0.1 mM EDTA, 0.5 mM spermidine, 0.15 mM spermine, 7 mM 2-mercaptoethanol) supplemented with complete protease inhibitor cocktail (Roche). Buffer C (350 mM NaCl, 10 mM HEPES pH8.0, 25% glycerol, 0.1 mM EDTA, 0.5 mM spermidine, 0.15 mM spermine, 7 mM 2-mercaptoethanol) supplemented with complete protease inhibitor cocktail (Roche) was used to isolate nuclear extracts. Protein concentration was quantified using Bradford Protein Assay Kit (Thermo Scientific). Lysates, nuclear and cytoplasmic fractions were separated on 10% SDS-PAGE, and transferred onto PVDF membrane using Towbin's transfer buffer. Membranes were blocked with 5% dry milk and 5% FBS in TTBS and incubated with primary antibodies diluted in blocking buffer. Primary anti-human OCT4 antibodies used were rabbit polyclonal ab19857 derived against the C-terminus of human OCT4, rabbit polyclonal ab18976 derived against amino acids 1–14 (both from Abcam); and mouse monoclonal sc-5279 (Santa Cruz Biotechnology) developed against the N-terminus. The anti- α -tubulin antibody (Sigma) was used as a loading control. Appropriate secondary HRP-conjugated antibodies were used and signal detected by enhanced chemiluminescence (Amersham, BioRad) on Kodak MM4000.

Immunocytochemistry

The same three anti-OCT4 antibodies were used for ICC: sc-5279 (Santa Cruz Biotechnology), ab19857, and ab18976 (both from Abcam). The following secondary antibodies were used: anti-mouse AlexaFluor568 and anti-rabbit AlexaFluor568. Alexa-Fluor488-conjugated phalloidin (Invitrogen) was used to stain actin filaments and DAPI to stain DNA. Images were obtained

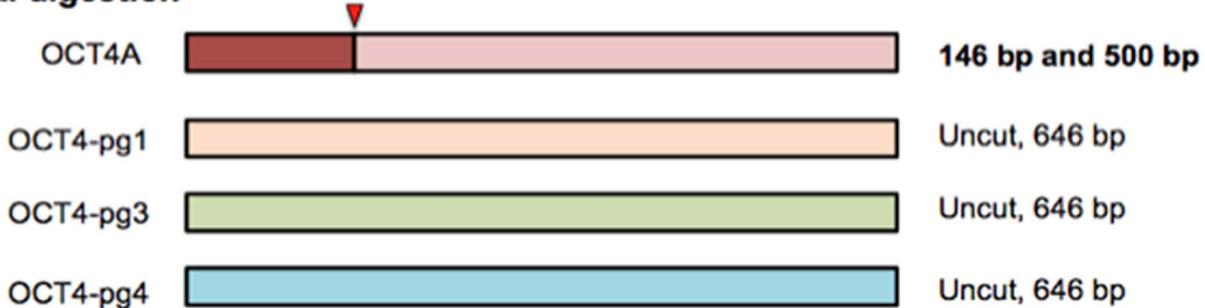
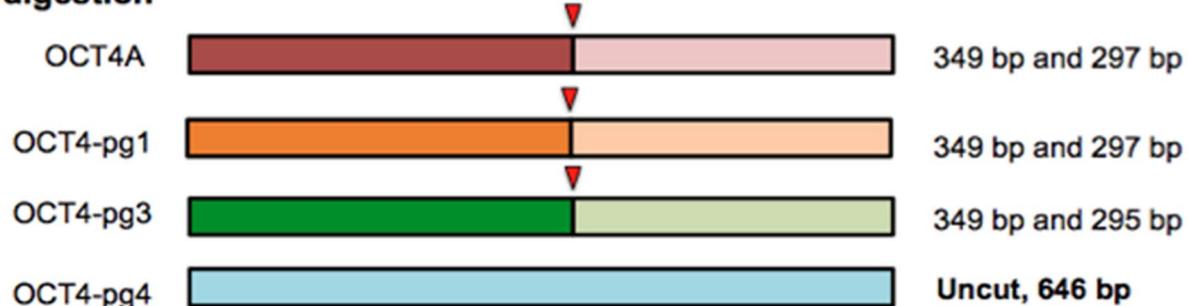
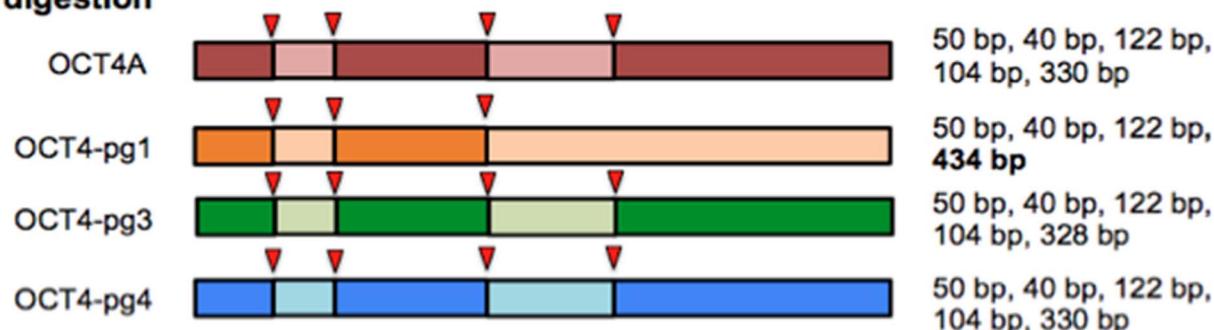
Apal digestion**Xhol digestion****BgII digestion****HinfI digestion**

Figure 2. Schematic representation of restriction sites in 646 bp OCT4 RT-PCR amplicon. Specific restriction sites were used to distinguish between embryonic OCT4A transcript and different pseudogenes. Red arrows show restriction sites. **Apal** restriction site is present only in embryonic OCT4A and can be used to distinguish embryonic form from all six pseudogenes; after restriction, a 146 bp and 500 bp long fragments are produced. **HinfI** digestion results in several smaller fragments among which the 434 bp fragment is specific only for OCT4-pg1. **BgII** digests only OCT4-pg3 into two fragments of 412 bp and 232 bp. **Xhol** does not digest OCT4-pg4.

doi:10.1371/journal.pone.0089546.g002

with appropriate excitation/emission on Leica confocal microscope using identical image acquisition settings.

Results

The Pool of OCT4 Transcripts in Various Cell Types is Composed of Messages Originating from its Gene and its Pseudogenes

To establish a baseline expression levels for stem cell genes in hESCs and iPSCs, we first examined expression of OCT4, NANOG and CRIPTO1 by RT-PCR. As expected, strong expression of all three stem cell genes was maintained during hESC (Figure 3A) and iPSC maintenance culture (data not shown). Starting 10 days after induction of differentiation, the expression of these embryonic markers gradually decreased and was undetectable for CRIPTO1 on day 14 and NANOG on day 17. Low levels of OCT4 transcript, however persisted and were still detectable on day 21 (Figure 3B). Expression of the same genes in control adult human fibroblasts resembled the expression of these genes in day 21 differentiated hESCs. We next examined whether expression of these genes can be detected in adult human fibroblasts cultured in 2% oxygen culture in the presence of FGF2. RT-PCR amplified transcripts could be detected for all three genes after 3 days and increased on day 7 (Figure 3C), confirming our previously published observations [37].

In addition to recognizing the OCT4A, our OCT4 primers were designed to amplify a fragment including 5' region which is common for OCT4A and pseudogenes 1, 3 and 4, but which is missing in all other pseudogene transcripts and OCT4B transcripts (Figure 1). The expected 646 bp amplicon, therefore could contain any of the four transcripts. This allowed us to examine the OCT4 amplicon in these cells in more detail. To identify the composition of the transcripts that could contribute to the 646 bp OCT4 amplicon at any time and to distinguish between the individual transcripts, we searched otherwise highly homologous sequences of OCT4 and OCT4-pg1, 3 and 4 for potentially unique endonuclease restriction sites. In addition to previously described ApaI site, that is specific only for OCT4A [55], we were able to identify three restriction sites that were then selected for differentiation between pseudogene transcripts (Figure 2). Restriction analysis of the amplicon from hESCs during differentiation is presented in Figure 3D. ApaI restriction analysis of the 646 bp PCR product identified the persistence of OCT4A until day 14 of differentiation, as evident by ApaI restriction of the amplicon into 146 and 500 bp fragments. By day 17 the amplicon could not be digested by ApaI any longer. This indicated that the embryonic OCT4A transcript was no longer present and the entire 646 bp amplicon on day 17 and after consisted of pseudogene transcripts. Digestion of the amplicon with XhoI, which digests all but the OCT4-pg4 indicated very low amounts of this transcript in hESCs and started gradually increasing between days 10 and 14 after differentiation. Digestion with BglII which cuts only OCT4-pg3 (412 and 232 bp fragments) revealed a gradual increase of this transcript 14 days after differentiation. Lastly, digestion with Hinfl which cuts all transcripts into 5 fragments except OCT4-pg1 (4 fragments) similarly indicated the start of OCT4-pg1 transcription on day 14 by the virtue of the longer, 424 bp fragment unique to OCT4-pg1. Overall, the amount of all pseudogene transcripts increased with a concomitant decrease in transcription of OCT4A over the 21 days of hESC differentiation. The pattern of OCT4A and pseudogenes expression in control fibroblasts resembled the pattern observed in differentiated hESCs and control fibroblasts, which expressed only OCT4 pseudogenes (Figure 3E). However, ApaI digestion of the OCT4 amplicon from fibroblasts grown in

2% oxygen and with 4 ng/ml FGF2— the conditions we have previously used to induce a phenotype switch to a regeneration competent cell [38] - identified embryonic OCT4A transcript at both day 3 and day 7 of culture in these conditions, indicating a reversal of expression from OCT4 pseudogenes to embryonic OCT4A (Figure 3G), resembling the pattern of ApaI digestion in hESC and iPSCs (Figure 3D and 3F). When a variety of other cell types were examined, only hESCs and teratocarcinoma cell lines (NTERA and NCCIT) expressed ApaI-sensitive OCT4A transcript (Figure 3H). Similar downregulation of OCT4A transcript was observed during differentiation of teratocarcinoma cells (Figure 4), however, simultaneous expression of pseudogenes started increasing much earlier during differentiation, and could be observed from day 5 onward.

To confirm our restriction digestion results, the 646 bp PCR amplicons from hESCs, 2%O₂/FGF2 treated fibroblasts and control fibroblasts were cloned into TOPO vector and three clones from each sample for each pseudogene and six clones from 2%O₂/FGF2 treated fibroblasts were sequenced. Sequencing confirmed our restriction data; hESCs and 2%O₂/FGF2 treated fibroblasts expressed only embryonic OCT4A, while control fibroblasts expressed OCT4-pg1, pg3 and pg4 (Tables S2–S7).

Translation and Localization of Protein Products of OCT4 Gene and its Pseudogenes

To determine whether any of the OCT4 transcripts were, in fact, translated into detectable proteins, Western blotting was performed. As demonstrated in Figure 5A, the OCT4 sc-5279 antibody, directed against N-terminal protein sequence, recognized several bands in teratocarcinoma cells (NCCIT) and in fibroblasts, regardless of culture conditions. The doublet that migrated at 45/43 kDa (OCT4A or OCT4pg1) and a band migrating around 37 kDa (likely somatic OCT4B) were present in all samples, while the two faster migrating proteins (30 and 20 kDa) were present only in fibroblasts grown in ambient oxygen, regardless off FGF2 supplementation. The proteins were detected in both cytoplasmic and nuclear fractions of fibroblasts. Knowing that FGF2/2%O₂ fibroblasts express OCT4A, however, it is reasonable to assume that some of the slower migrating doublet represents embryonic OCT4A. In control fibroblasts these bands are likely contributed by pseudogenes 1, 3 and 4, as the antibody should recognize sequences in the N-terminal of these proteins (Figure 5A). The second antibody, ab19857 (against C-terminal), similarly recognized the doublet of 45/43 kDa and a protein migrating around 37 kDa, but failed to recognize any other protein products. The amount of detected protein at 37 kDa was significantly higher in cells grown in low oxygen and potentiated by addition of FGF2 (Figure 5B). The third antibody ab18976, derived against the N-terminal OCT4 sequence could, in principle detect OCT4A and pseudogenes 1 and 3. Intensity of the band in treated fibroblast was significantly lower than in hESCs (Figure 5C). A 31 kDa band was detected in both control and treated fibroblasts and could represent OCT4pg4. The identity of the 26 kDa band detected with this antibody remains unknown. An additional 19 kDa protein, that could be the product of OCT4pg3, was recognized in control fibroblasts, but was absent in stem cells and treated fibroblasts.

Immunocytochemistry revealed OCT4 presence in cell nuclei of hESCs regardless of the antibody used, as expected (Figure 6). The N-terminal specific antibody sc-5279 recognized weak nuclear signal in 2%O₂/FGF2 fibroblasts. The C-terminal ab19857 recognized nuclear localized protein in hESCs, and nuclear and cytoplasmic proteins in control and treated fibroblasts. As none of the OCT4B variants localize to the nucleus, the nuclear signal can

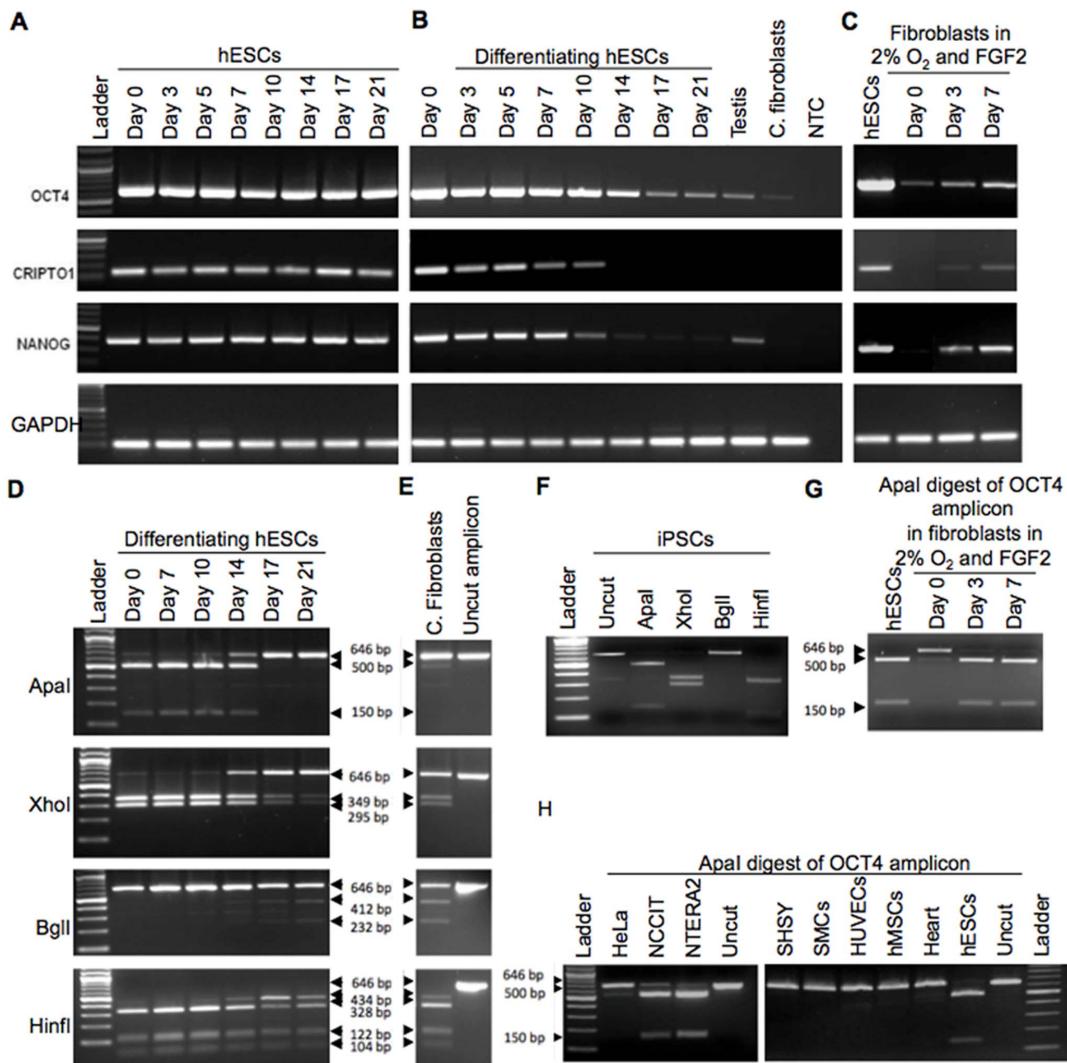


Figure 3. RT-PCR for expression of embryonic stem cell specific genes. OCT4, NANOG and CRIPTO1 in human ESCs during culture (A), during 21 days after induction of hESC differentiation (B), and in adult human dermal fibroblasts cultured in 2% oxygen with FGF2 supplementation (C). Restriction digest of 646 bp OCT4 amplicon from human embryonic stem cells (D), control fibroblasts (E), and iPSCs (F). Apal restriction digest of OCT4 amplicon in fibroblasts grown in 2% oxygen and FGF2 supplementation (G), and various transformed, multipotent and differentiated cells (H). NCCIT – teratocarcinoma, NTERA2– teratocarcinoma, SHSY – neuroblastoma, SMCs – smooth muscle cells, HUVECs – human umbilical vein endothelial cells, hMSCs – human mesenchymal stem cells. Only the embryonic OCT4A contains Apal restriction site.

doi:10.1371/journal.pone.0089546.g003

be contributed only by OCT4A or pseudogene OCT4-pg1 derived protein [48,51]. Considering the individual cell type transcript profile, it is possible that in treated fibroblasts the antibody recognized OCT4A. The OCT4-pg4 protein has yet to be detected by any anti-OCT4 antibody, and OCT4-pg3 localizes to the cytoplasm [51]. The sc-8629, another C-terminal specific antibody, recognized nuclear OCT4A in hESCs and cytoplasmic proteins in both control and treated fibroblasts. Cytoplasmic staining in both fibroblasts may be identifying OCT4A, OCT4B, OCT4pg1 [42], and/or OCT4-pg3 [51]. These data corroborate our previously published results [35] and indicate that expression of OCT4 in somatic cells is under the control of an oxygen and FGF2-sensitive mechanism.

Discussion

OCT4 has received considerable attention during the past years and its expression has been reported in various cell types and across a spectrum of differentiation cell states. Due to the possibility of multiple transcript variants and protein isoforms, as well as expressed pseudogenes being detected, more rigorous analyses of OCT4 are required [44,46]. Existence of high number of pseudogenes for OCT4 is not an exception. Excessive frequency of pseudogenes for embryonic stem cells-specific genes was also described for Nanog (10 pseudogenes), Stella (16 pseudogenes), and Cripto-1 (6 pseudogenes) [47,52].

We identified expression of Oct-4pg1, Oct-4pg3 and Oct-4pg4 pseudogenes in normal adult human fibroblasts using PCR based approach followed by pseudogene-specific restriction digestion and sequencing. Our study confirms that expression of OCT4A protein is exclusive to hESCs and germ cell-derived tumor cell

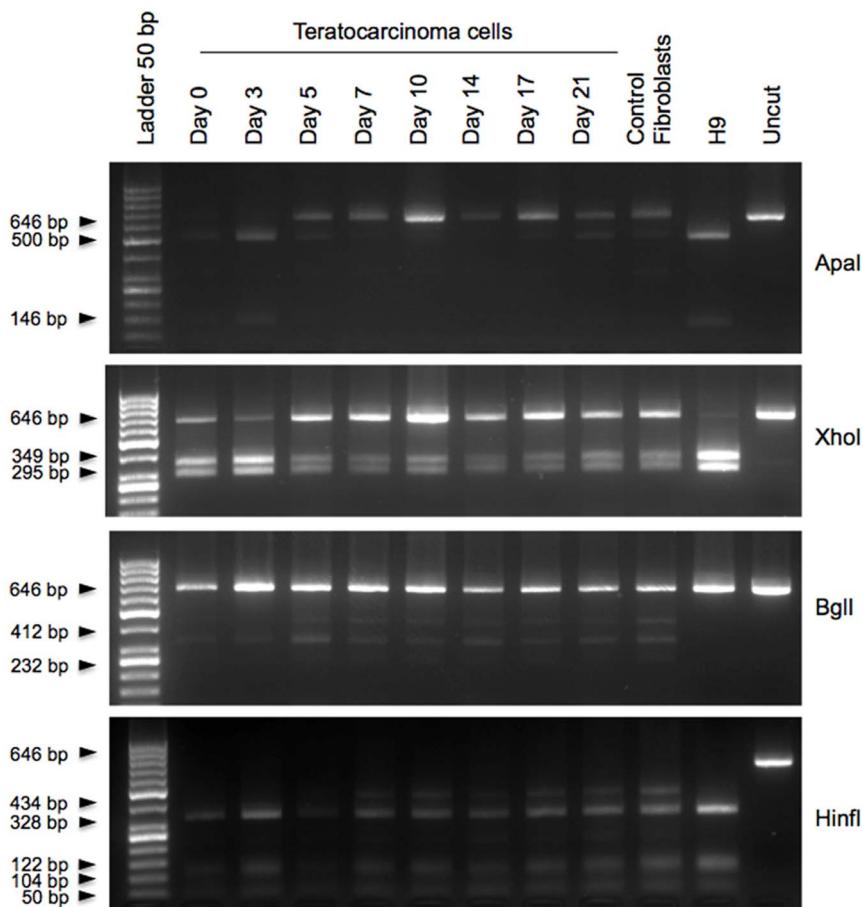


Figure 4. Restriction digest of 646 bp OCT4 amplicon from human embryonic carcinoma cells during 21 days of differentiation in medium containing retinoic acid.

doi:10.1371/journal.pone.0089546.g004

lines. None of the other examined somatic cells, somatic tumor cell lines or adult stem cells expressed OCT4A. In an exact opposite pattern, all cell types examined except hESCs and germ cell tumor-derived cell lines expressed pseudogenes 1, 3 and 4. Monitoring the relationship between OCT4A expression and expression of pseudogenes in hESCs during differentiation, an interesting pattern emerges; a gradual decrease of OCT4A occurs in a parallel with an increase in expression of pseudogenes. This may indicate a tight regulatory relationship between OCT4A and its pseudogenes, even though the OCT4 gene and all the pseudogenes are located on different chromosomes [47]. One of the pseudogenes, namely OCT4-pg3 is located within the 445 kb of NANOG, STELLA and GDF3 locus on chromosome 12 [56], significance of which still needs to be examined.

It is interesting that similar changes occur in reverse during acquisition of regeneration competence in adult human fibroblasts. After a short term culture in reduced oxygen and supplementation with FGF2 [35,38] human fibroblast cease expressing OCT4 pseudogenes in favor of OCT4A. We have reported previously that these cells do not acquire the pluripotent phenotype [35], but rather undergo a phenotype switch that demonstrates regeneration competence *in vivo* [38]. Cells demonstrate an increased life span, doubled population doubling capacity, maintain a normal karyotype and do not form tumors in SCID mice. They do, however, demonstrate long term engraftment and acquire PAX7 positive phenotype when transplanted into a skeletal muscle defect

in mice [38]. Significance and relevance of OCT4A expression, if any, during acquisition of regeneration competence, is not clear and is currently under investigation. Nevertheless, this is the first study in which a switch from pseudogene to OCT4A expression has been reported and induced by selective manipulation of culture conditions. We speculate that pseudogene silencing is induced in adult somatic cells using our culture system, an event that may be required for re-activation of OCT4A. While clearly not pluripotent, regeneration-competent cells express low levels of OCT4A. It has been shown that OCT4 is controlling pluripotency in a quantitative manner [57]. Specifically, high level of OCT4 expression drives ES cells to endoderm and mesoderm lineages, while stem cells with low level of OCT4 differentiate into trophoectoderm. Only a specific level of OCT4 can maintain stem cells in a pluripotent state. These observations suggest that OCT4 is different from many known transcription regulators that appear to function in an on-off manner. Consequently, low levels of embryonic OCT4A may not be functionally relevant for either induction or maintenance of pluripotency, but may be relevant for induction of this novel regeneration-competent cell phenotype. Our data indicate that chromatin plasticity may be an inherent feature of adult fibroblasts and that the manifestation of this plasticity may be influenced by the extracellular environment. Our data also support the notion that regeneration competence does not depend on cell pluripotency and may be regulated by a novel, OCT4A-dependent mechanism. OCT4 gene ablation in stem cell

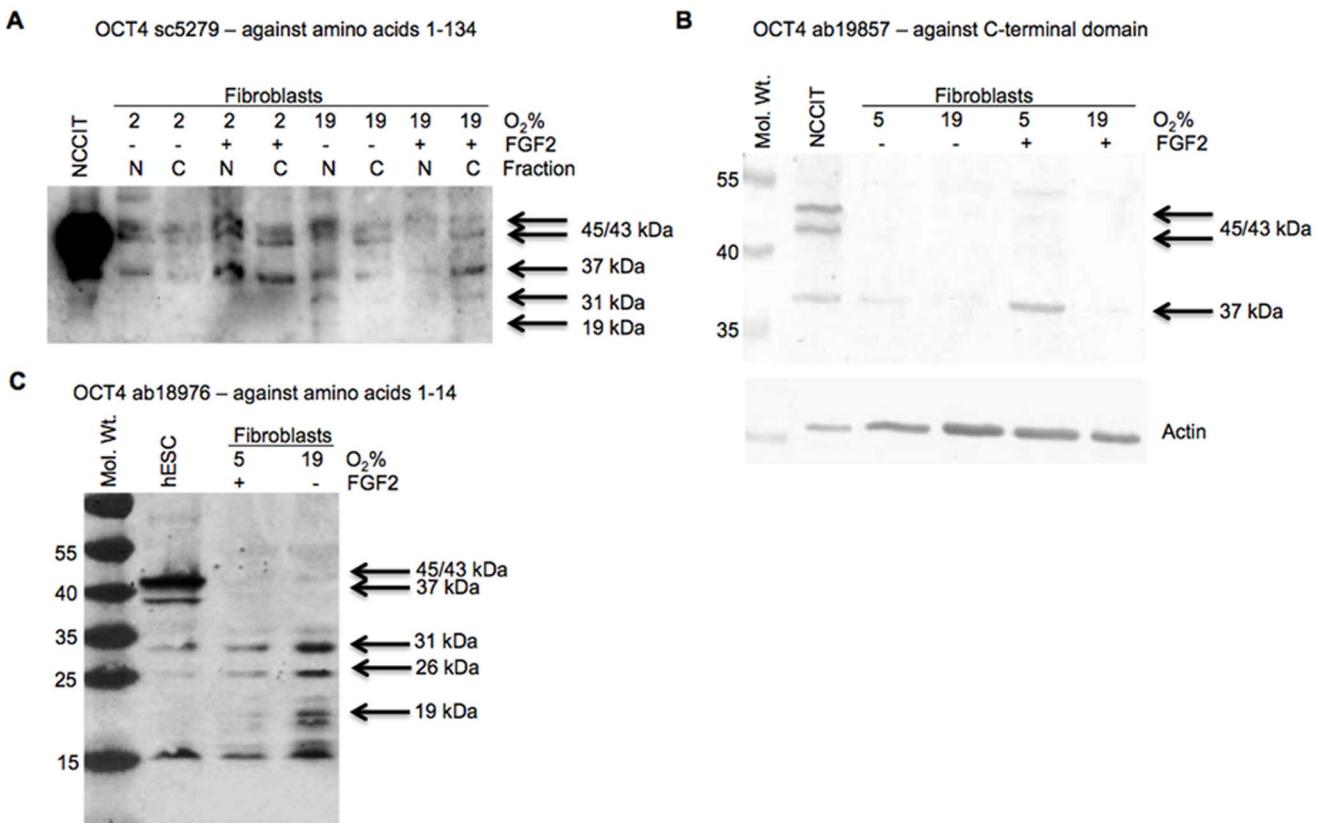


Figure 5. Protein expression of stem cell specific genes. Western blots for OCT4 in human embryonic stem cells (hESC), teratocarcinoma cells and adult human dermal fibroblasts (A, C) or muscle fibroblasts (B) using different antibodies. Complete cell lysates or nuclear and cytoplasmic fractions (N and C, respectively) were separated on 10% SDS PAGE, transferred to PVDF and blotted with different antibodies. Fibroblasts were grown for 7 days under low (2 or 5%) or ambient (19%) oxygen, in the presence or absence of FGF2 (+ and -, respectively).

compartments of several somatic tissues, including intestinal epithelium, bone marrow (hematopoietic and mesenchymal lineages), hair follicle, brain and liver in mice revealed no abnormalities in homeostasis or regenerative capacity indicating that OCT4 is dispensable for both self-renewal and maintenance of somatic stem cells in adult mammal. Neither OCT4 nor its

pseudogenes, for example, are required for multipotency of human stromal cells from different tissues [49]. There is no common regulatory network that has been shown to exist between the various somatic stem cell types and that can be compared to gene regulatory network that maintains pluripotency in ES cells with OCT4 as its cornerstone [52,58]. It is more likely that

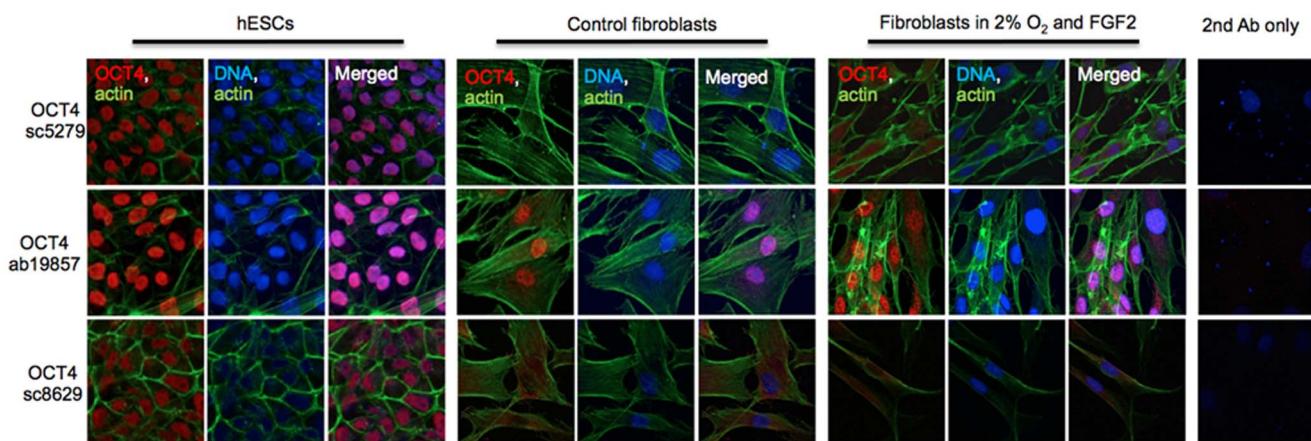


Figure 6. Immunocytochemistry of OCT4 in human embryonic stem cells (hESCs) and fibroblasts. Cells were cultured in control (ambient 19% O₂, no FGF2) and 2% oxygen with FGF2 supplementation for 7 days. Cells were stained with three different OCT4 antibodies, as indicated. Red – OCT4, Green – actin; Blue – DNA.

doi:10.1371/journal.pone.0089546.g006

developmental potency of tissue-specific stem cells is governed by intrinsic and extrinsic signals from the microenvironment in which they exist [59].

Several reports argue that detection of OCT4A in somatic cells is an artifact; possibly a consequence of DNA contamination and/or inability of currently available methods to distinguish between OCT4A and its pseudogenes. Misidentification of OCT4 in somatic tissues and cancer cell lines can also result from detection of pseudogene transcripts by RT-PCR and/or nuclear detection of pseudogene proteins by ICC. Most studies indeed do not exclude the possibilities of pseudogene expression [16,44,46,59] and do not distinguish between OCT4 isoforms [46]. All OCT4 protein isoforms are identical in their C-terminal region and distinct in N-terminal region so the location of the epitope recognized by any anti-OCT4 antibody should be considered first [42,46] and combined with a diligent analysis of the OCT4 transcript. Adding the possibility that OCT4 pseudogenes translate into proteins and that OCT4-pg1 can localize to the nucleus [51] and confer transcriptional activation of some reporter gene constructs [49], diligent analysis should be employed. This may likely require a combination of experimental approaches, such as RT-PCR, restriction analysis for various pseudogenes, and protein analysis. Additional functional studies, such as chromatin immunoprecipitation, electromobility shift assay and/or transcription analysis of downstream genes may need to be included to reach a valid conclusion.

Besides accurately characterizing OCT4 expression, relevance of OCT4 expression should be considered for each chosen experimental system. While OCT4A expression may be critically important when studying pluripotency and self-renewal, examining expression of OCT4A in studies involving stem cells from adult compartments or in somatic cell cancers may be of little functional relevance [59,60].

Supporting Information

Figure S1 Differences between Oct4 pseudogenes based on the coding sequences. N-terminal domains from all pseudogenes are similar to embryonic Oct4.

(TIF)

Table S1 OCT4 and its pseudogenes. Source: <http://www.ncbi.nlm.nih.gov/sites/entrez>.

(DOCX)

Table S2 CLUSTAL 2.1 multiple sequence alignment of Oct4A, Oct4pg1, Oct4pg3 and Oct4pg4 sequences. Oct4 primers: F, 5'- GTTGATCCTCGGACCTGGCTA-3'; R, 5'-

References

- Hansis C, Tang YX, Grifo JA, Krey LC (2001) Analysis of Oct-4 expression and ploidy in individual human blastomeres. Mol Hum Reprod 7(2): 155–161.
- Mitalipov SM, Kuo H-C, Hennebold JD, Wolf DP (2003) Oct-4 expression in pluripotent cells of the rhesus monkey. Biol Reprod 69(6): 1785–1792.
- Schöler HR, Balling R, Hatzopoulos AK, Suzuki N, Gruss P (1989) Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. The Embo Journal 8(9): 2551–2557.
- Cauffman G, Van de Velde H, Liebaers I, Van Steirteghem A (2005) Oct-4 mRNA and protein expression during human preimplantation development. Mol Hum Reprod 11(3): 173–181.
- Cauffman G, Liebaers I, Van Steirteghem A, Van de Velde H (2006) POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos. Stem Cells 24(12): 2685–2691.
- Hansis C, Grifo JA, Krey LC (2000) Oct-4 expression in inner cell mass and trophectoderm of human blastocysts. Mol Hum Reprod 6(11): 999–1004.
- Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, et al. (2007) Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. Stem Cells 25(2): 500–510.
- Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122(6): 947–956.
- Gerrard L, Zhao D, Clark AJ, Cui W (2005) Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. Stem Cells 23(1): 124–133.
- Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, et al. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95(3): 379–391.
- Pesce M, Gross MK, Schöler HR (1998) In line with our ancestors: Oct-4 and the mammalian germ. Bioessays 20(9): 722–732.
- Schöler HR, Hatzopoulos AK, Balling R, Suzuki N, Gruss P (1989) A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. The EMBO Journal 8(9): 2543–2550.
- Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L, et al. (2004) Oct4 is required for primordial germ cell survival. EMBO Reports 5(11): 1078–1083.
- Palumbo C, van Roozenendaal K, Gillis AJ, van Gurp RH, de Munink H, et al. (2002) Expression of the PDGF alpha-receptor 1.5 kb transcript, OCT-4, and c-

GGTTGCCTCTCACTCGGTTCT-3'; (reverse complement: AGAACCGAGTGAGAGGGCAACC). Amplicon length is 646 bp. Primer sequences are shown in yellow, **bold** is the sequence amplified with these primers. Purple are polymorphisms in primer annealing sites. Nucleotide polymorphisms exploited for enzyme specific digestion are marked in red and the whole enzyme recognition site in green.

(DOCX)

Table S3 Alignment of the cloned 646 bp amplicon amplified from embryonic stem cells (H9) – clone 1, 8 and 9– aligned to Oct4A mRNA sequence from GenBank (NM_002701.4).

(DOCX)

Table S4 Alignment of the 646 bp amplicon amplified from fibroblasts grown in 2%O₂ and FGF2 (CRL2352 treated) – colony 1, 3, 4, 5, 6 and 7– aligned to Oct4A mRNA sequence from GenBank (NM_002701.4).

(DOCX)

Table S5 Alignment of the 646 bp amplicon amplified from control fibroblasts (CRL2352 untreated) – colony 12, 14 and 28– aligned to Oct4pg1 mRNA sequence from GenBank (NR_002304.2).

(DOCX)

Table S6 Alignment of the 646 bp amplicon amplified from control fibroblasts (CRL2352 untreated) – colony 3, 7 and 11– aligned to Oct4pg3 mRNA sequence from GenBank (NR_036440.1).

(DOCX)

Table S7 Alignment of the 646 bp amplicon amplified from control fibroblasts (CRL2352 untreated) – colony 2 and 5– aligned to Oct4pg4 mRNA sequence from GenBank (NR_034180.1).

(DOCX)

Acknowledgments

We gratefully acknowledge help by Marsha Rolle, George Pins and Glenn Gaudette with experimental protocols. We would like to thank Vicky Huntress for help with confocal imaging.

Author Contributions

Conceived and designed the experiments: MJ SA. Performed the experiments: MJ OK AG CM LV. Analyzed the data: MJ. Wrote the paper: MJ CM PR TD.

- KIT in human normal and malignant tissues. Implications for the early diagnosis of testicular germ cell tumours and for our understanding of regulatory mechanisms. *J Pathol* 196(4): 467–477.
15. Cheng L, Sung M-T, Cossu-Rocca P, Jones TD, MacLennan GT, et al. (2007) OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. *J Pathol* 211(1): 1–9.
 16. Looijenga LHJ, Stoop H, de Leeuw HPJC, de Gouveia Brazao CA, Gillis AJM, et al. (2003) POU5F1 (OCT3/4) identifies cells with pluripotent potential in human germ cell tumors. *Cancer Res* 63(9): 2244–2250.
 17. Sperger JM, Chen X, Draper JS, Antosiewicz JE, Choi CH, et al. (2003) Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. *Proc Natl Acad Sci U.S.A.* 100(23): 13350–13355.
 18. de Jong J, Looijenga LHJ (2006) Stem cell marker OCT3/4 in tumor biology and germ cell tumor diagnostics: history and future. *Crit Rev Oncogenesis* 12(3–4): 171–203.
 19. Takahashi K, Tanabe K, Ohnuki M, Narita M, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5): 861–872.
 20. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318(5858): 1917–1920.
 21. Boyer LA, Mathur D, Jaenisch R (2006) Molecular control of pluripotency. *Curr Opin Genetics Dev* 16(5): 455–462.
 22. Greber B, Lehrach H, Adjaye J (2007) Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs to support hESC self-renewal. *Stem Cells* 25(2): 455–464.
 23. Kucia M, Halasa M, Wysoczynski M, Baskiewicz-Masiuk M, Moldenhawer S, et al. (2007) Morphological and molecular characterization of novel population of CXCR4+ SSEA-4+ Oct-4+ very small embryonic-like cells purified from human cord blood: preliminary report. *Leukemia* 21(2): 297–303.
 24. Sovalar H, Scrofani M, Eidschenk A, Pasquet S, Rimelen V, et al. (2011) Identification and isolation from either adult human bone marrow or G-CSF-mobilized peripheral blood of CD34(+) /CD133(+) /CXCR4(+) /Lin(–)CD45(–) cells, featuring morphological, molecular, and phenotypic characteristics of very small embryonic-like (VSEL). *Exp Hematol* 39(4): 495–505.
 25. Kucia MJ, Wysoczynski M, Wu W, Zuba-Surma EK, Ratajczak J, et al. (2008) Evidence that very small embryonic-like stem cells are mobilized into peripheral blood. *Stem Cells* 26(8): 2083–2092.
 26. Ratajczak MZ, Machalinski B, Wojakowski W, Ratajczak J, Kucia M (2007) A hypothesis for an embryonic origin of pluripotent Oct-4(+) stem cells in adult bone marrow and other tissues. *Leukemia* 21(5): 860–867.
 27. Roche S, Richard M-J, Favrot M-C (2007) Oct-4, Rex-1, and Gata-4 expression in human MSC increase the differentiation efficiency but not hTERT expression. *J Cell Biochem* 101(2): 271–280.
 28. Pacini S, Carnicelli V, Trombi L, Montali M, Fazzi R, et al. (2010) Constitutive expression of pluripotency-associated genes in mesodermal progenitor cells (MPCs). *PLOS ONE* 5(3): e9861.
 29. Tai M-H, Chang C-C, Kiupel M, Webster JD, Olson LK, et al. (2005) Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. *Carcinogenesis* 26(2): 495–502.
 30. Atsasi Y, Mowla SJ, Ziaeem SAM, Bahrami A-R (2007) OCT-4, an embryonic stem cell marker, is highly expressed in bladder cancer. *Int J Cancer* 120(7): 1598–1602.
 31. Chiou S-H, Yu C-C, Huang C-Y, Lin S-C, Liu C-J (2008) Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin Cancer Res* 14(13): 4085–4095.
 32. Ezech UI, Turek PJ, Reijo RA, Clark AT (2005) Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma. *Cancer* 104(10): 2255–2265.
 33. Monk M, Holding C (2001) Human embryonic genes re-expressed in cancer cells. *Oncogene* 20(56): 8085–8091.
 34. S Zangrossi, Marabese M, Broggini M, Giordano R, D'Erasmo M, et al. (2007) Oct-4 expression in adult human differentiated cells challenges its role as a pure stem cell marker. *Stem Cells* 25(7): 1675–1680.
 35. Page RL, Ambady S, Holmes WF, Vilner L, Kole D, et al. (2009) Induction of stem cell gene expression in adult human fibroblasts without transgenes. *Cloning Stem Cells* 11(3): 417–426.
 36. Bhartiya D, Shaikh A, Nagvenkar P, Kasiviswanathan S, Pethe P, et al. (2011) Very Small Embryonic-like Stem Cells with Maximum Regenerative Potential get Discarded during Cord Blood Banking and Bone Marrow Processing for Autologous Stem Cell Therapy. *Stem Cells Dev*.
 37. Kucia M, Reca R, Campbell FR, Zuba-Surma E, Majka M, et al. (2006) A population of very small embryonic-like (VSEL) CXCR4(+)/SSEA-1(+)/Oct-4+ stem cells identified in adult bone marrow. *Leukemia* 20(5): 857–869.
 38. Page RL, Malcuic C, Vilner L, Vojtic I, Shaw S, et al. (2011) Restoration of skeletal muscle defects with adult human cells delivered on fibrin microthreads. *Tissue Eng Part A*; 17(21–22): 2629–40.
 39. Kashpur O, Lapointe D, Ambady S, Ryder EF, Dominko T (2013) FGF2-induced effects on transcriptome associated with regeneration competence in adult human fibroblasts. *BMC genomics*; 14: 656.
 40. Takeda J, Seino S, Bell GI (1992) Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic Acids Res* 20(17): 4613–4620.
 41. Atsasi Y, Mowla SJ, Ziaeem SAM, Gokhale PJ, Andrews PW (2008) OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. *Stem Cells* 26(12): 3068–3074.
 42. Wang X, Zhao Y, Xiao Z, Chen B, Wei Z, et al. (2009) Alternative translation of OCT4 by an internal ribosome entry site and its novel function in stress response. *Stem Cells* 27(6): 1265–1275.
 43. Lee J, Kim HK, Rho J-Y, Han Y-M, Kim J (2006) The human OCT-4 isoforms differ in their ability to confer self-renewal. *J Biol Chem* 281(44): 33554–33565.
 44. Liedtke S, Stephan M, Kögl G (2008) Oct4 expression revisited: potential pitfalls for data misinterpretation in stem cell research. *Biol Chem* 389(7): 845–850.
 45. Wang X, Dai J (2010) Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem Cells* 28(5): 885–893.
 46. Liedtke S, Enczmann J, Waclawczyk S, Wernet P, Kögl G (2007) Oct4 and its pseudogenes confuse stem cell research. *Cell Stem Cell* 1(4): 364–366.
 47. Pain D, Chirn G-W, Strassel C, Kemp DM (2005) Multiple retropseudogenes from pluripotent cell-specific gene expression indicates a potential signature for novel gene identification. *J Biol Chem* 280(8): 6265–6268.
 48. Panagopoulos I, Möller E, Collin A, Mertens F (2008). The POU5F1P1 pseudogene encodes a putative protein similar to POU5F1 isoform 1. *Oncol Rep* 20(5): 1029–1033.
 49. Kaltz N, Funari A, Hippauf S, Delorme B, Noël D, et al. (2008) In vivo osteoprogenitor potency of human stromal cells from different tissues does not correlate with expression of POU5F1 or its pseudogenes. *Stem Cells* 26(9): 2419–2424.
 50. Suo G, Han J, Wang X, Zhang J, Zhao Y, et al. (2005) Oct4 pseudogenes are transcribed in cancers. *Biochem Biophys Res Commun* 337(4): 1047–1051.
 51. Zhao S, Yuan Q, Hao H, Guo Y, Liu S, et al. (2011) Expression of OCT4 pseudogenes in human tumours: lessons from glioma and breast carcinoma. *J Pathol* 223(5): 672–682.
 52. Redshaw Z, Strain AJ (2010) Human hematopoietic stem cells express Oct4 pseudogenes and lack the ability to initiate Oct4 promoter-driven gene expression. *J Negat Results Biomed* 9(1): 2.
 53. Korneev SA, Park JH, O'Shea M (1999) Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene. *J Neurosci* 19(18): 7711–7720.
 54. Yano Y, Saito R, Yoshida N, Yoshiki A, Wynshaw-Boris A, et al. (2004) A new role for expressed pseudogenes as ncRNA: regulation of mRNA stability of its homologous coding gene. *Journal Mol Med* 82(7): 414–422.
 55. Panagopoulos I, Möller E, Isaksson M, Mertens F (2008) A PCR/restriction digestion assay for the detection of the transcript variants 1 and 2 of POU5F1. *Genes Chromosome Canc* 47(6): 521–529.
 56. Clark AT, Rodriguez RT, Bodnar MS, Abeyta MJ, Cedars MI, et al. (2004) Human STELLAR, NANOG, and GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. *Stem Cells* 22(2): 169–179.
 57. Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet* 24(4): 372–376.
 58. Berg JS, Goodell MA (2007) An argument against a role for Oct4 in somatic stem cells. *Cell Stem Cell* 1(4): 359–360.
 59. Lengner CJ, Welstead GG, Jaenisch R (2008) The pluripotency regulator Oct4: a role in somatic stem cells? *Cell Cycle* 7(6): 725–728.
 60. Cantz T, Key G, Bleidissel M, Gentile L, Han WD, et al. (2008) Absence of OCT4 expression in somatic tumor cell lines. *Stem Cells* 26(3): 692–697.