

Research article

Open Access

Promoter analysis of the rabbit *POU5F1* gene and its expression in preimplantation stage embryos

Julianna Kobolak¹, Katalin Kiss^{1,2}, Zsuzsanna Polgar³, Solomon Mamo^{1,4}, Claire Rogel-Gaillard⁵, Zsuzsanna Tancos^{1,6}, Istvan Bock^{6,7}, Arpad G Baji¹, Krisztina Tar⁷, Melinda K Pirity⁷ and Andras Dinnyes*^{1,6}

Address: ¹Micromanipulation and Genetic Reprogramming Group, Agricultural Biotechnology Center, Szent-Györgyi A. u. 4. H-2100 Gödöllő, Hungary, ²National Medical Center Cell Biology Department; Daróci u. 24. H-1113 Budapest, Hungary, ³Faculty of Natural Sciences, Constantine the Philosopher University, Slovakia, ⁴University College Dublin, Lyons Research Farm, Newcastle Co. Dublin, Ireland, ⁵INRA, UMR1313, Laboratoire de Génétique Animale et Biologie Intégrative, 78350 Jouy en Josas, France, ⁶Molecular Animal Biotechnology Laboratory, Szent Istvan University, Pater K. u. 1. H-2103 Gödöllő, Hungary and ⁷BioTalentum Ltd., Aulich Lajos u. 26. H-2100 Gödöllő, Hungary

Email: Julianna Kobolak - kobolak@abc.hu; Katalin Kiss - katalin.kiss@kkk.org.hu; Zsuzsanna Polgar - polgarzs@abc.hu; Solomon Mamo - msolomon@abc.hu; Claire Rogel-Gaillard - claire.rogel-gaillard@jouy.inra.fr; Zsuzsanna Tancos - tancoszs@abc.hu; Istvan Bock - istvan.bock@biotalentum.hu; Arpad G Baji - baji@abc.hu; Krisztina Tar - krisztina.tar@biotalentum.hu; Melinda K Pirity - melinda.pirity@biotalentum.hu; Andras Dinnyes* - andrasdinnyes@yahoo.com

* Corresponding author

Published: 4 September 2009

Received: 12 December 2008

BMC Molecular Biology 2009, 10:88 doi:10.1186/1471-2199-10-88

Accepted: 4 September 2009

This article is available from: <http://www.biomedcentral.com/1471-2199/10/88>

© 2009 Kobolak et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The *POU5F1* gene encodes the octamer-binding transcription factor-4 (Oct4). It is crucial in the regulation of pluripotency during embryonic development and widely used as molecular marker of embryonic stem cells (ESCs). The objective of this study was to identify and to analyse the promoter region of rabbit *POU5F1* gene; furthermore to examine its expression pattern in preimplantation stage rabbit embryos.

Results: The upstream region of rabbit *POU5F1* was subcloned sequenced and four highly conserved promoter regions (CRI-4) were identified. The highest degree of similarity on sequence level was found among the conserved domains between rabbit and human. Among the enhancers the proximal enhancer region (PE-1A) exhibited the highest degree of homology (96.4%). Furthermore, the CR4 regulator domain containing the distal enhancer (DE-2A) was responsible for stem cell-specific expression. Also, BAC library screen revealed the existence of a processed pseudogene of rabbit *POU5F1*. The results of quantitative real-time PCR experiments showed that *POU5F1* mRNA was abundantly present in oocytes and zygotes, but it was gradually reduced until the activation of the embryonic genome, thereafter a continuous increase in *POU5F1* mRNA level was observed until blastocyst stage. By using the XYClone laser system the inner cell mass (ICM) and trophoblast portions of embryos were microdissected and examined separately and *POU5F1* mRNA was detected in both cell types.

Conclusion: In this study we provide a comparative sequence analysis of the regulatory region of rabbit *POU5F1* gene. Our data suggest that the *POU5F1* gene is strictly regulated during early mammalian development. We proposed that the well conserved CR4 region containing the DE-2A enhancer is responsible for the highly conserved ESC specific gene expression. Notably, we are the first to report that the rabbit *POU5F1* is not restricted to ICM cells only, but it is expressed in trophoblast cells as well. This information may be well applicable to investigate further the possible phylogenetic role and the regulation of *POU5F1* gene.

Background

The *POU5F1* gene belongs to the POU (Pit-Oct-Unc) family of transcription factors, that encodes the octamer-binding transcription factor-4 (Oct4) [1]. In mouse, before the zygotic gene activation the active *POU5F1* mRNA is present in the oocyte. The zygotic expression is activated around the 4-cell stage and later restricted to the pluripotent cells of the inner cell mass (ICM) and to the epiblast. Following implantation, the expression is down-regulated and limited to the primordial germ cells (PGC) but silenced in all somatic cells [2]. *In vitro* embryo studies in bovine, pig, rhesus monkey and in human have shown that the protein is present in the trophoblast cells of blastocyst stage embryos and it is not restricted to the pluripotent ICM cells [3-6]. *In vitro*, *POU5F1* is highly expressed in human and mouse ESCs. These cells lose their pluripotency during differentiation, therefore the *POU5F1* expression is diminished [7,8].

The *POU5F1* protein is among the core group of transcription factors that induces and controls stemness in ESCs. It sustains pluripotency through feed-forward and feedback transcriptional mechanism and it also plays a crucial role in the early mammalian development [7,9].

The expression of *POU5F1* is controlled by cis-regulatory elements, located 5' upstream from the initiation site of transcription [2,10]. The regulatory region of *POU5F1* is highly conserved among species; and usually contains four conserved regions (CR) within the promoter. The TATA-less minimal promoter (MP) region is always located within the first conserved region (CR1) of the upstream sequence of the gene. This minimal promoter contains further primary regulatory elements, such as Sp1/Sp3, and hormone responsive element (HRE) binding sites. Reporter gene expression experiments in mouse with LacZ revealed that two elements, the proximal enhancer (PE) and the distal enhancer (DE) are essential in the cell-specific regulation of *POU5F1*. The proximal enhancer (PE) is located about 1.2 kb upstream, within conserved regions (CR2 and CR3) and is responsible for *POU5F1* expression in embryonic ectoderm and mouse embryonal carcinoma (EC) cells. Finally, the distal enhancer (DE) located about 2 kb upstream, also within a conserved region (CR4) and drives the *POU5F1* expression in the morula, ICM, ESC, embryonic germ (EG) and PGC cells of the mouse [2,11,12].

Rabbit (*Oryctolagus cuniculus*) is a classical experimental animal model due to its physiological and immunological properties; it is preferentially used in pulmonary, cardiovascular and metabolic studies, as well as for antibody production and drug screening. In mouse, a wide range of genetic methods have been developed so far, and the successful application of the same technology in rabbits would be very desirable. Recent publications about gener-

ating rabbit ESCs [13,14] and successful somatic cell nuclear transfer (SCNT) experiments [15,16] raise the possibility for increasing the use of rabbit model systems for wider applications.

Although the rabbit *POU5F1* cDNA and the genomic DNA sequences were recently published [17], the isolation and the detailed characterisation of its regulatory region is still missing. Moreover, its expression during embryo preimplantation has not yet been well characterized [18]. This study for the first time describes isolation and sequencing of the rabbit orthologue of *POU5F1* 5' regulatory region and the identification of its phylogenetically conserved regions. We also examined the function of the individual truncated promoter regions using GFP reporter gene system. In addition we also examined the expression pattern of rabbit *POU5F1* in individual preimplantation stage embryos using quantitative real-time PCR.

Results

Isolation of the rabbit *POU5F1* upstream region

Recently, the MHC region of rabbit was mapped [19] and partial *POU5F1*-like gene has been sequenced on BAC clones LBAB-841A3 and LBAB-85810. Based on multi-species sequence comparison (human, mouse and bovine *POU5F1* genes), primers were designed for the conserved regions (oct4-435 and oct4-186, Table 1), and applied for PCR screens and further sequencing.

The sequenced 7.5 kb *POU5F1*-like region was confirmed as the rabbit's entire *POU5F1* gene by multiple sequence comparison. Five highly conserved regions have been found, which perfectly matched with the well-characterized five exons of the reference sequences. The lengths of the sequences of these conserved, putative exonic regions were quite similar but not identical, indicating the existence of non-conserved intronic regions. The complete rabbit *POU5F1* sequence is now referred to as [EF194086](#) and mapped to the MHC region, 12q1.1 [17].

To confirm the coding sequence of *POU5F1*, RACE PCR was applied on blastocyst samples. The gene-specific inner primers were bound to the conserved region of the gene; the outer primers were annealed to the 5'UTR and the polyA tail. The expressed and coding region was found to be 1083 bp, which matched exactly with the previously predicted exons. The rabbit cDNA sequence was identical to human, mouse and bovine in 84%, 83% and 87%, respectively. The cDNA sequence is now referred as [EF062856](#)[17].

Sequence analysis of the rabbit *POU5F1* promoter region

Following confirmation of the *POU5F1*-like sequence as a rabbit coding sequence, the 2.2 kb fragment of the *POU5F1* upstream region was sequenced and aligned to

Table 1: Primers used in BAC library screen, promoter analysis and real-time RT-PCRs

Primer name	Primer (5'-3') sequence	Position	Primer size (bp)	Tm (°C)
oct4-435-F	AGCTTAGCTTCAAGAACATG	+3665	20	60.0
oct4-435-R	AGGAGTACAGTGCAGTGAAG	+4704	20	60.0
oct4-186-F	AGCAGAAACCCTCGTGCAGG	+3759	20	60.0
oct4-186-R	TCTGGCGCCGGTTACAGAAC	+4181	20	60.0
*oct4-pseudo-F	GCTAAACAGAAAGAAGTTTGCC	* +420	22	57.0
*oct4-pseudo-R	GAACAGTCACTGCTTGATCGTTT	* +870	23	58.1
oct4-cDNA-F	GCTCTACAGAAAGAACTCGAGCAG	+3640	24	57.9
oct4-cDNA-R	CGAGTACAGGGTAGCAAAGTGAG	+4869	23	60.0
CR1-ATGmut	TCTATGGGGGAAGGAGGGCG	+5	20	60.0
CR1	AGGCTGGTGGCATAAAACAC	-558	20	60.0
CR1+(CR2+CR3)	GCCAGACTAGAGCCCAACAG	-1585	20	60.0
CR1+(CR2+CR3)+CR4	GGGAAATTGTGGAGGAGGAC	-2030	20	60.7
CR4	TCTGTCTGCTTGGTGGTGTCT	-1670	20	59.9
Rabbit-oct4-RT-F	CGAGTGAGAGGCAACTTGG	+4418	19	55.5
Rabbit-oct4-RT-R	CGGTTACAGAACCACACACG	+4730	20	56.1
*Rabbit-oct4p-RT-F	CTAAACAGAAAGAAGTTTGCC	* +421	21	53.7
*Rabbit-oct4p-RT-R	CGCAGCTTACACATGTACT	* +594	19	52.3
Mouse-oct4-RT-F	ATGCCGTGAAGTTGGAGAAG	+343	20	56.5
Mouse-oct4-RT-R	GGTCTGGCTGAACACCTTTC	+3309	20	55.9

The initial ATG is signed as +1, the positions show the genomic locations of the primer. F: forward, R: reverse * pseudogene-specific primers, the positions show the pseudogene location of the primers

its orthologues (Figure 1). The alignment proved the presence of four extensively conserved regions (CR1-4) and three enhancer regions in the rabbit *POU5F1* 5' upstream sequence. These CR regions are known to be the functional regions of the *POU5F1* promoter in several species [11,20]. The different regions in rabbit showed extensive identity compared to the human, bovine, dog and to mouse sequences (Figure 1). In rabbit, the TATA-less minimal promoter (MP) that is located within the CR1 region, extends until -221 bp, and contains both the overlapping binding sites of Sp1/Sp3 (-120/-110), the hormone responsive element (-111/-94) and three G/C rich regions (Figure 1). The E-box sequence in rabbit (CACTTG), that plays a regulatory role in the control of transcription by binding basic helix-loop-helix (bHLH) transcription factors, is located within the proximal enhancer Region 1B (PE-1B;-945/-940). The proximal enhancer Region 1B (-968/-940) is localized within the CR2 region (-1023/-824), while the proximal enhancer Region 1A (PE-1A;-1232/-1192) is located outside of the CR2 region (-1480/-1376). This PE-1A region exceptionally shows less similarity to all reference species. The CR3 region is localized upstream of PE-1A sequence. The distal enhancer region (DE-2A) is located within the CR4 region (-2450/-1901), last from -2025 to -2004 bp. Transcription factor (TF) binding site, Oct4/Sox2 is located (-1977/-1962) in the distal enhancer, within the CR4 region. We have identified several G/C-rich motifs present in all of the five mammalian sequences; however the one near the PE-1A

enhancer we found a CCCACCC motif present only in rabbit. This motif partially overlaps the CCCTCCC sequence. The schematic draw of organization of rabbit *POU5F1* promoter region is shown in Figure 2.

Table 2 shows the pairwise comparison of the rabbit and four other mammalian *POU5F1* regulatory regions. The analysis unambiguously revealed that the human has the closest potential evolutionary orthologue of the rabbit *POU5F1* regulatory region. This is supported with the highest homology that could be observed between rabbit and human in all of the four conserved regions and in the PE-1B and DE-2A enhancers. The complete minimal promoter (0/-221 bp) and PE-1A enhancer were the two exceptional regions where the dog sequence showed the highest homology compared to the rabbit sequence (Table 2).

Functional analysis of the promoter regions

Four different reporter constructs were electroporated into mouse ESCs to test the functionality of the predicted conserved region in the *POU5F1* promoter (Figure 3). The minimal (proximal) promoter (CR1), the proximal promoter with proximal enhancers (CR1+(CR2 and CR3), the proximal promoter with distal enhancer (CR1+CR4) and the entire promoter region (CR1+(CR2 and CR3)+CR4) were tested. The fluorescent activity of the minimal promoter (CR1) was very weak compared to the control ubiquitin promoter and to the basic auto-fluores-

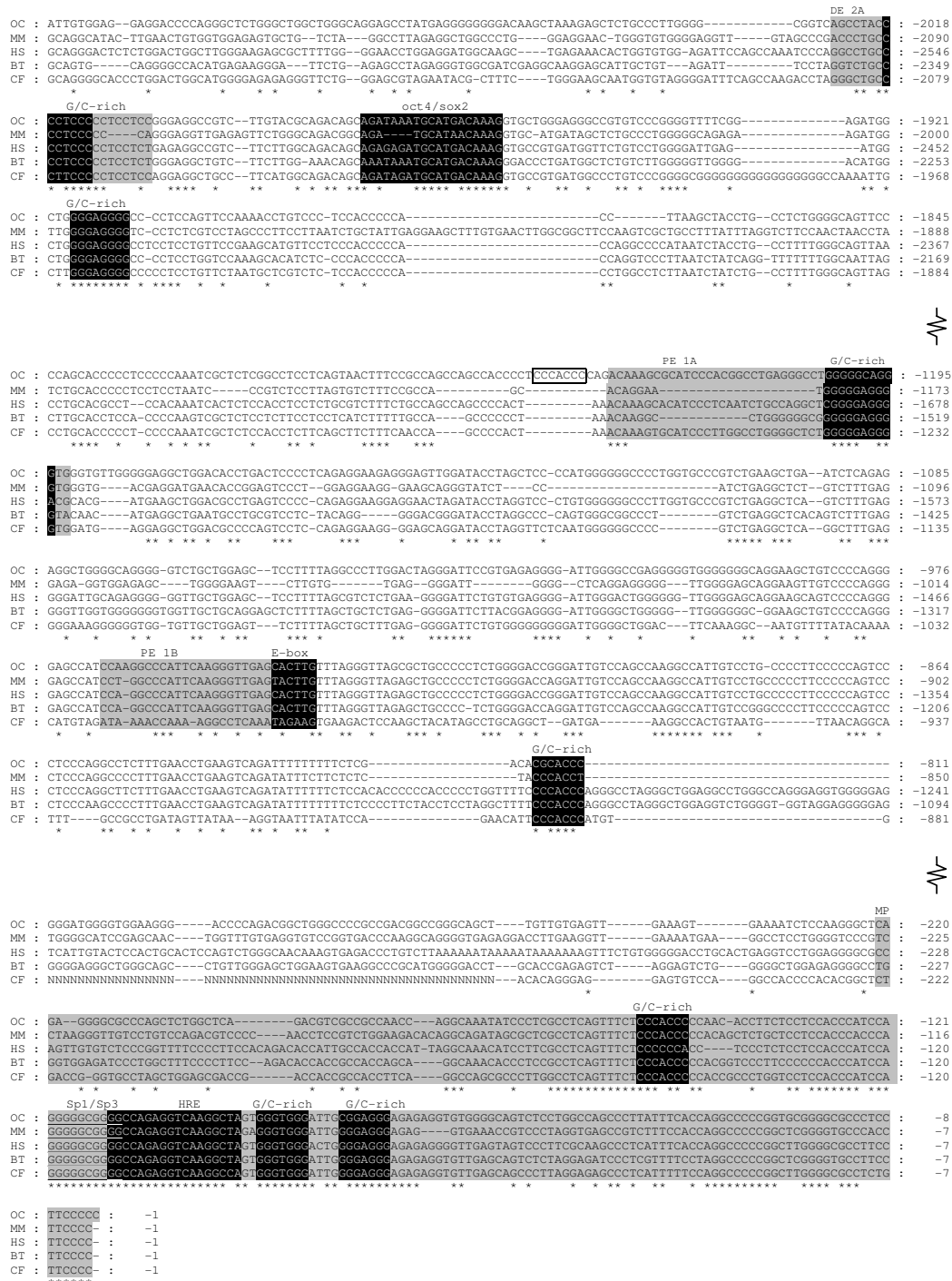


Figure 1
Alignments of the rabbit (OC) POU5F1 regulatory regions with its mammalian orthologues. The region of distal enhancer, site 2A (DE 2A); proximal enhancer, sites 1A and 1B (PE 1A, PE 1B) and the minimal promoter (MP) are shaded in grey colour. The highly conserved GC-rich motifs like GGG(A/T)GGG, CCC(A/T)CCC and the putative transcription factor binding sites are shaded in black colour. A CCCACCC motif near PE 1A present only in rabbit is rimmed; this motif partially overlaps a CCCTCCC motif (bolded). The Sp family binding site is underlined. Nucleotides have been numbered relative to the translation start site of the rabbit POU5F1 gene. Oc, *Oryctolagus cuniculus*; Mm, *Mus musculus*; Hs, *Homo sapiens*; Bt, *Bos taurus*, and Cf, *Canis familiaris*.

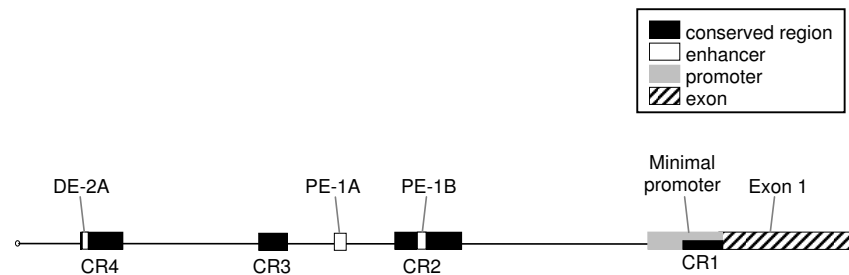


Figure 2
Organization of the *POU5F1* promoter region in rabbit (2641 bp).

cence of non-transfected control R1 ESCs (Figure 3). Therefore, the rabbit *POU5F1* regulatory region is capable to express in mouse ESC as a minimal promoter. The addition of CR2+CR3 regions did not increase the level of expression substantially compared to the CR1 fragment. However, when the CR4 region was co-transfected with the minimal promoter (CR1), the level of expression was increased more than two fold. Similarly, the expression was also two fold higher compared to that of the CR1 alone, when the full promoter region [CR1+(CR2 and CR3)+CR4] was transfected or the region of CR1+(CR2 and CR3) was co-transfected with the CR4. However, the intensity of the expression was significantly less compared to the ubiquitin promoter (B6U-3 ESC) or to the control pmaxGFP vector expression (Figure 3). When its expression was compared to the mouse *POU5F1* minimal promoter (CR1) and distal enhancer-containing promoter region construct (Oct4-GiP) [21], the later exhibited a lower expression intensity in average (Figure 3).

***POU5F1* expression in preimplantation stage embryos**

The expression of rabbit *POU5F1* gene was analyzed by quantitative real-time RT-PCR in preimplantation stage

Table 2: The homology (%) between the nucleotide sequences of conserved regions and the complete upstream region

Conserved region of <i>Oc</i>	<i>Mm</i>	<i>Hs</i>	<i>Bt</i>	<i>Cf</i>
CR1	82.3	88.4	86.8	84.8
CR2	92.0	93.9	90.7	-
CR3	79.6	80.9	73.8	68.2
CR4	67.6	83.6	78.0	67.1
PE 1A	76.5	73.2	77.8	83.3
PE 1B	96.0	96.4	96.4	-
DE 2A	87.5	95.0	90.0	85.7
MP	74.2	76.8	79.1	79.3
5' region	52.5	57.7	57.0	51.9

Pairwise comparison method was used to compare the *POU5F1* regulatory regions of the rabbit and of the four mammalian species. The largest homologies of the regions are bolded; *Oc*, *Oryctolagus cuniculus*; *Mm*, *Mus musculus*; *Hs*, *Homo sapiens*; *Bt*, *Bos taurus*, and *Cf*, *Canis familiaris*; (-) no data.

rabbit embryos, and compared with the mouse *POU5F1* expression. In all examined stages the mRNA of *POU5F1* was detectable in the embryos with various intensities. The mRNA was present in oocytes and zygotes at a higher level (Figure 4a and 4b), but the levels continuously decreased until the embryonic genome activation, which occurs between late 8- to 16-cell stages during rabbit pre-implantation development [18,22,23]). Following the embryonic genome activation, the expression levels increased continuously until the blastocyst stage (Figure 4b). The *POU5F1* transcript was also quantified in the isolated ICM and trophoblast portions of the rabbit blastocysts and the relative levels were compared taking the blastocyst transcript as calibrator. The comparison revealed that the *POU5F1* gene was clearly expressed in both cell types but at a significantly higher ($p < 0.05$) level in the ICM compared to the trophoblast cells (Figure 4c).

Identification of pseudogenes

BAC library screen was performed with primers oct4-435 and oct4-186 (Table 1) and BAC clones LBAB-304A07 and LBAB-779H10 were identified. Both clones contain a *POU5F1*-like sequence, and although they lack introns, they contain remnants of polyA tail flanked by direct repeats, all of the characteristics of retrotransposed pseudogenes [24,25]. Based on sequencing the flanking region, the two BAC clones were found to be identical, containing the same genomic region. Yet, the chromosomal localization of these BAC clones is unknown.

The cDNA of *POU5F1* gene and its pseudogene share a 99% overall sequence identity [see Additional file 1]. Sequence analysis revealed that there is an eight nucleotide deletion in the pseudogene in position 796 bp from the initial ATG that causes a frameshift in the sequence. This results the encoding of a 162 amino acid-long truncated protein instead of the 360 amino acid-long functional protein. This truncated form lacks the functional POU domain, but still has a 92% similarity on amino acid level until the frameshift caused STOP codon. The sequence of the pseudogene is referred as [EU191070](#).

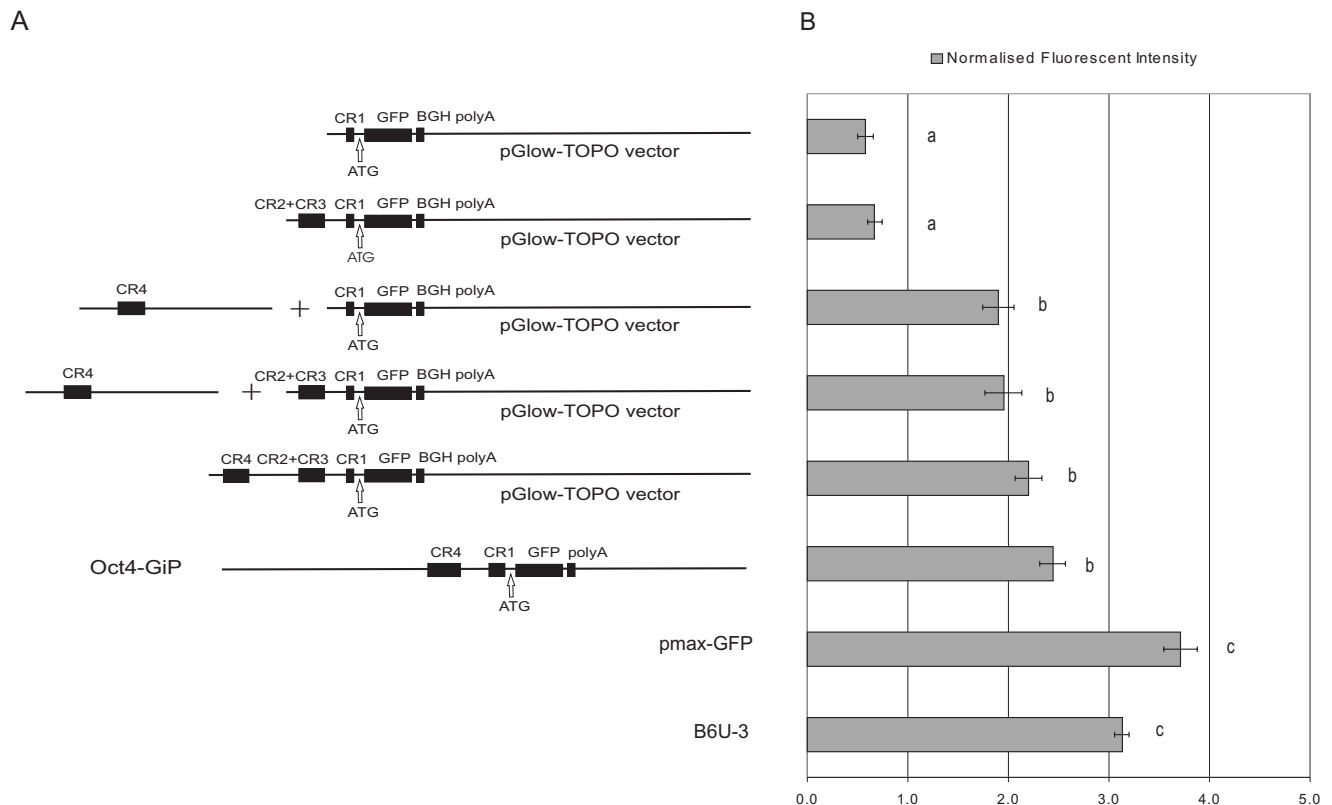


Figure 3

Expressional analysis of rabbit *POU5F1* promoter region in mouse ES cells. A. Schematic view of different constructs used in the assay. B. Fluorescent activity was measured 72 hours post-nucleofection. Data was normalised with the auto-fluorescence of untreated RI ESCs. CR1, the minimal promoter showed a weak level of basic expression in the mouse ESCs. When the proximal enhancer fragment (CR2 and CR3) was part of the regulator region the expression showed no significant change. However, when the vector containing the distal enhancer region (CR4) was co-transfected either with the minimal promoter (CR1) or with the minimal promoter and proximal enhancers (CR1+(CR2 and CR3)), the expression level increased significantly. While the entire regulatory region (CR1+(CR2 and CR3)+CR4) was transfected, the expression peaked to its highest level. If compared to the Oct4-GiP cells containing the mouse CR4+CR1 region, no significant differences were found [20]. S.E. \pm values are marked on each column. Different letters mark significant differences between values ($P < 0.05$).

Furthermore, to test whether the pseudogene was retrotransposed back near to a functioning promoter, we examined its expression using pseudogene-specific primers in RT-qPCR experiment. The analysis showed that the pseudogene is not expressing at any preimplantation developmental-stages in rabbit [see Additional file 1].

Discussion

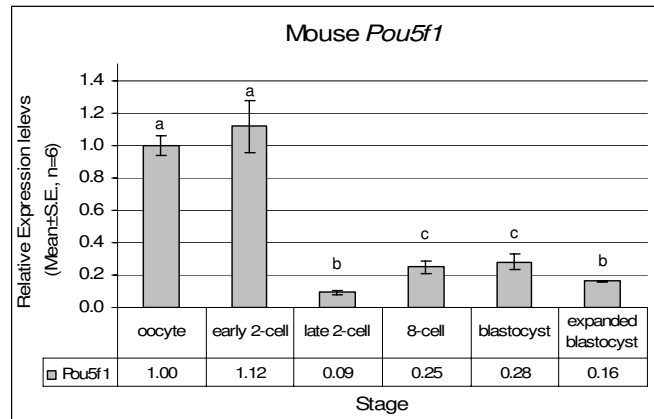
This is the first study describing the cloning and the analysis of the complete regulatory region of the rabbit *POU5F1* gene. We have also cloned the rabbit *Pou5f1* cDNA and analyzed the expression pattern in preimplantation-stage embryos. In addition, we identified a processed pseudogene, having a high degree of similarity.

Our results demonstrate that both the gene structure and the sequence of *POU5F1* are extremely conserved among

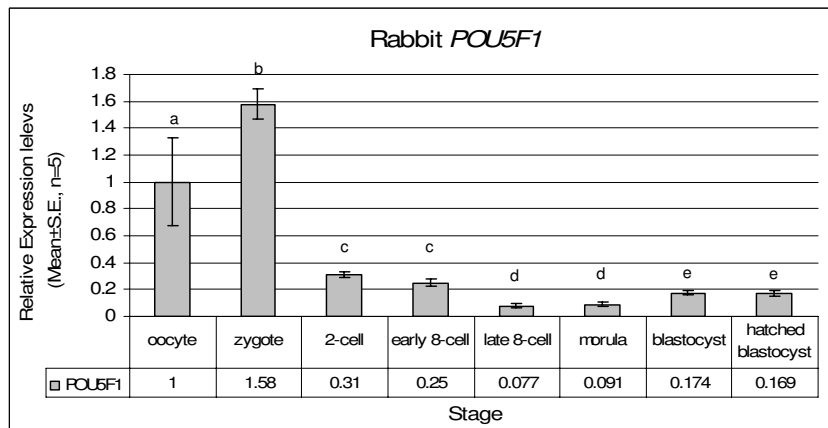
various species (rabbit and its murine, human, bovine and dog counterparts). As in other species, *POU5F1* was mapped to the MHC region, and contained 5 exons with a longer first intron and 3' UTR. The putative protein sequence is almost identical with that of other species. This indicates the highly conserved role of *POU5F1* in early vertebrate development.

In our study, the 2.2 kb fragment of the rabbit *POU5F1* upstream region was isolated and aligned to its orthologues (Figure 1; GenBank: [EU189135](#)). Comparing the 5' region of human, bovine, mouse and dog *POU5F1* sequences to the rabbit, all the functionally important regions could be unquestionably defined within 3 kb upstream of its transcription start site. In comparison to the mouse, the rabbit sequence lacks large intersections between CR1 and CR2 regions, but the functional regions

A



B



C

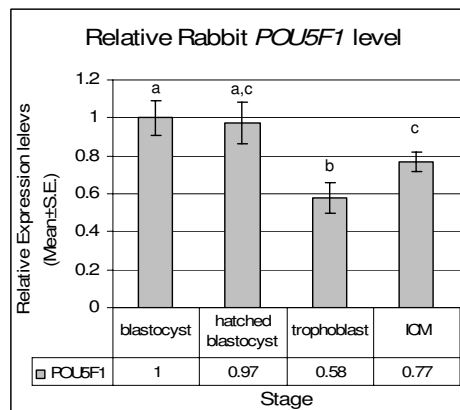


Figure 4

Oct4 expression in preimplantation stage rabbit embryos. A. Quantitative real time RT-PCR analysis of *POU5F1* expression in mouse and B. rabbit preimplantation stage embryos. C. Relative expression level of *POU5F1* in ICM and trophoblast portion of rabbit blastocyst. The mean values of gene expression of six individual samples and \pm S.E. are shown on the diagrams. For comparison of expression through developmental stages (B), the oocyte level was taken as calibrator, while for comparison of TE and ICM samples (C), the blastocyst level was taken as calibrator. Note the scale of the y-axis differs in the case of the two species. Different letters mark significant differences between values ($P < 0.05$).

in all the four species were almost identical. Several conserved repetitive sequences, E-boxes, the binding site of Pem, Pax-3 and Sp1 transcription factors [26] were identified. Furthermore, the HRE binding site was also identified which was shown previously to be important in RA-induced *POU5F1* down regulation. The newly identified CCC(A/T)CCC motifs that we have identified only in rabbit might gain importance when studying the differences of *POU5F1* gene regulation in mammalian species.

The CR4 region is of particular interest for pluripotent stem cell research [21], because this region is responsible for the *POU5F1* expression in ESCs [2]. Previously, the extended comparative sequence analysis of other mammals (pig, dog, rat) clearly showed three highly consensus sites inside the CR4 region [27]. In rabbit, all three sequences are identical; this suggests the functional significance and a highly conserved regulatory role in mammals. Using comparative sequence analysis we observed that the rabbit and the human *POU5F1* regulatory regions share the greatest homology among the investigated four species.

Our functional assessment studies on mouse ESCs verified the sequence analysis results. In our experiments the rabbit CR4 region containing the promoter fragments showed the highest expression level in mouse ESCs. This suggests that the rabbit CR4 promoter region confer the same transcriptional regulation activity as the mouse CR4 sequence in mouse ESCs. Also, the rabbit construct showed no significant difference in expression intensity compared to the mouse *POU5F1* promoter (Oct4-GiP).

The quantitative real-time RT-PCR showed that there is a high level of maternal origin *POU5F1* mRNA in oocytes and zygotes, which decreased gradually until 8 to 16-cell stage. The transiently reduced transcript levels increased continuously after the embryonic genome activation (Figure 4b). These data are in agreement with the measured (Figure 4a) and published gene expression patterns of mouse [28]. However, since the embryo genome activation in rabbit starts later than in mouse, a difference in the timing of endogenous *POU5F1* expressions was detected between mouse and rabbit.

In mouse, the expression becomes confined to the ICM following the morula stage, up-regulated in the primitive ectoderm, and eventually becomes confined to primordial germ cells [2]. In rabbit, we found that the *POU5F1* was expressed in both the trophoblast and the ICM at the mRNA level. In bovine, porcine, rhesus monkey and human preimplantation stage embryos the *POU5F1* protein was also localized in both cell types [3-6]. In bovine the presence of *POU5F1* protein in trophoblast cells might act as a suppressor of extraembryonic lineage-spe-

cific gene expression during the preimplantation development which allows an extensive proliferation of the trophoblast before the implantation [4,29]. These results indicate that the restriction of expression in blastocyst stage is species-specific [3].

Finally, we have identified a processed, retrotransposed pseudogene of the rabbit *POU5F1*. Retrotransposed genes are generated from mRNA and they integrate back to the genome after reverse transcription. Due to this process, these kinds of pseudogenes usually lack a functioning promoter, or endure some modification, which make them unable to code the functional or full-length proteins. Here we confirmed that the newly identified rabbit *POU5F1* pseudogene was not expressed during rabbit preimplantation development. Further studies will be needed to clarify its expression in further developmental stages and in differentiated tissues as some pseudogenes are known to possess certain regulatory roles [30-32]. Bioinformatics analysis revealed that retrotransposition frequency of ESC-specific genes appears far to exceed that of non-ESC specific genes, and multiple highly homologous pseudogenes likely exist for all of the ESC-specific genes [33,34]. This is a characteristic of human [35] and also of mouse [1,36,37]. It is not clear yet whether this abundant presence of pluripotency pseudogenes in mouse and human would be also an indication of unknown functions. In human the *POU5F1P1* pseudogene encodes a putative protein similar to *POU5F1* isoform 1, which is localized close to the amplified region of a variety of human malignancies [38]. Recent reports on human *POU5F1* pseudogenes suggested - especially in somatic cells, tissue-specific stem cells, and non-germ cell tumours - that some re-evaluation of expression data on *POU5F1* might be needed due to the use of primers which cannot distinguish pseudogenes from the coding one [39-41]. In other species - including the mouse and rabbit - the functional relevance of pseudogenes of *POU5F1* is still uncovered.

Conclusion

This study provides the first comprehensive analysis of rabbit *POU5F1* regulatory domains. Detailed sequence analysis of the isolated rabbit *POU5F1* promoter revealed high sequence conservation in the identified enhancer domains. The strong conservation of regulatory domain sequences shows the importance of the control of *POU5F1* expression in early mammalian development. We have shown that the highest homology of enhancers can be identified between human and rabbit. This may increase the importance to use the rabbit as a comparative functional genomics tool for human medical research.

Methods

All chemicals, unless otherwise stated, were purchased from Invitrogen Inc. (Carlsbad, California, USA).

The animal experiments were executed in full compliance with European and Hungarian laws and regulations, and were approved by the Agricultural Biotechnology Center, Animal Experimentation Committee, Gödöllő, Hungary.

Rabbit BAC library

From a rabbit genomic BAC library [19] four BAC clones were isolated (LBAB-841A3; LBAB-85810; LBAB-304A07 and LBAB-779H10). The clones are stored at the INRA resource centre [42]. The library screening was performed as reported previously [19].

Cloning and sequencing

The obtained PCR fragments were cloned into pCR2.1-TOPO vector. BAC clones and plasmid DNAs were purified by QIAGEN Plasmid Maxi Kit and QIAprep Spin Miniprep Kit, respectively (QIAGEN). Sequencing was performed in the ABC Sequencing Laboratory with Applied Biosystems BigDye v3.1 Ready Reaction Terminator Cycle Sequencing Kit (ABI) using the ABI PRISM 3100 Genetic Analyser automated nucleotide sequencer.

The genomic sequence of *Oryctolagus cuniculus* 5' regulatory region of POU5F1 gene is deposited at the GenBank under accession number [EU189135].

Comparative genomics

The POU5F1 sequence was submitted and referred as [GenBank: EF062856] for cDNA, and as [GenBank: EF194086] for genomic sequence. For multiple alignments the following genomic, cDNA and protein sequences were used: [GenBank: NT_007592.14], [GenBank: NM_002701.4], [GenBank: NP_002692.2] for human; [GenBank: MGI_101893], [GenBank: NM_013633.2], [GenBank: NP_038661.2] for mouse; and [GenBank: NW_001494164.1], [GenBank: NM_174580.1], [GenBank: NP_777005.1] for bovine, respectively.

For the comparative analysis the POU5F1 5' upstream sequences of mouse [GenBank: NT_039649.7], human [GenBank: NT_007592.14g] bovine [GenBank: AF022986] and dog [GenBank: NC_006594.2] sequences were aligned by using ClustalW2 program [43]. Pairwise comparison of nucleotide sequences was performed on EMBOSS Pairwise Alignment Algorithms program [44].

RACE-PCR

Rapid Amplification of cDNA Ends (RACE) PCR was performed according to the manufacturer's instruction. As outer primers oct4-5UTR and oligo(dT) primers were used, while oct4-435 and oct4-186 primers were applied as inner, gene-specific primers (Table 1).

Sample materials

All samples were isolated from Hycote Hybrid rabbits. Oocytes were collected 16 hours post-hCG injection, while one-cell stage embryos were collected 18 hours post-hCG injection from adult female rabbits. Zygotes were cultivated in EBSS media [45] at 38.5°C under 5% CO₂ in air. After removing the *zona pellucida* with Pronase (Sigma), samples were collected in each developmental stage and kept frozen at -86°C until use. ICM and trophoblast samples from blastocyst were separated using the XYClone laser system (Hamilton Thorne Biosciences) [46]. Briefly, a blastocyst held by two large bore capillaries at the trophoblast region and at the ICM region, respectively, was gently pulled apart while exposed to laser pulses at the area of the trophoblast adjacent to the ICM. When the laser pulses cut through the cells, the two separated parts of the blastocyst were stored separately for further analyses. Due to the visually controlled nature of the procedure, the trophoblast sample is guaranteed to be ICM-cell free, however, the ICM sample might contain a small number of trophoblast cells. Mouse oocytes and *in vivo* embryo samples were collected as described earlier [47].

Functional analysis of the POU5F1 promoter

The conserved regions of the POU5F1 promoter were identified by multiple sequence alignments, amplified with PCR (Table 1) and subcloned into the pGlow TOPO® TA Expression Kit, containing the GFP reporter gene. The start codon of translation was mutated with PCR-based ATG-mutagenesis (Table 1).

The constructs were nucleofected into mouse R1 ESC by using the AMAXA Nucleofector System and Mouse ES cell Nucleofector Kit (Amaxa Biosystems). Electroporation was carried out based on the manufacturer's protocol using program no. A23. Briefly, 5 × 10⁶ cells were used for the nucleofection and about 5 µg of plasmid DNA were applied, after copy number normalization of the different constructs. Neomycin selection was started 24 hours post-nucleofection by applying 400 µg/ml G418 on the cultures. For transfection control the Amaxa Nucleofector Kit control vector pmaxGFP was used.

The R1 ESC line [48] (provided by Dr. Andras Nagy; Samuel Lunenfeld Research Institute, Toronto, Canada) was cultured using standard protocols [49]. Oct4-GiP transgenic mouse ESC line, expressing the EGFP reporter gene under the control of mouse POU5F1 [CR4+CR1] promoter region was kindly provided by Dr. Austin Smith (University of Edinburgh's Centre for Genome Research, Edinburgh, UK) [21]. The transgenic ubiquitin promoter-GFP expressing ESC line (B6U 3) was isolated from C57BL/6-Tg(UBC-GFP)30Scha/J transgenic mice (Stock

Number: 004353; The Jackson Laboratory) in our laboratory (unpublished data).

Fluorimetric assay

For GFP assays, cells were harvested at 72 hours post-transfection and the fluorescent intensity was measured using the Perkin-Elmer LS-50B Luminescence Spectrophotometer (PerkinElmer Life and Analytical Sciences Inc.; excitation: 375 nm, detection: 510 nm). Fluorescence was corrected for background activity shown by cells transfected with the pGlow-TOPO[®] vector alone. Cell number variation between individual experiments was corrected measuring the cell protein content (Bio-Rad DC Protein Assay system, BioRad Laboratories).

Primer design and quantitative real-time RT-PCR analysis

To design *POU5F1* pseudogene-specific primers, we initially performed a BLASTn [50] search to reveal any rabbit *POU5F1* pseudogene sequence up to date. This led to the only pseudogene sequence referred EU191070. Based on the sequence alignment of the *POU5F1* mRNA and its pseudogene [see Additional file 1], we designed a primer to the eight nucleotide deletion region, to avoid amplification of the *POU5F1* mRNA (Table 1).

The procedures of RNA isolation and cDNA synthesis were done as described in our earlier work [51]. Briefly, the mRNA was extracted from individual oocytes and embryos by Dynabeads[®] mRNA Direct Micro Kit according to the manufacturer's instructions. DNase treatment (Ambion-ABI) was applied on all RNA samples before cDNA synthesis. Genomic DNA was isolated from rabbit skin by phenol-chloroform extraction according to the standard methodology. Complementary DNA synthesis was performed with MMLV Reverse Transcriptase and oligodT primers, using the manufacturer's protocol.

Real-time PCR was performed in Rotor-Gene 3000 Real-time PCR Cycler (Corbett Research, Mortlake, Australia). The reaction mixture consisted of SYBR[®] Green JumpStart[™] Taq ReadyMix[™] (Sigma), 300 nM of each primer (Table 1), and cDNA of 0.10 embryo equivalent in a final volume of 25 μ l.

The reaction conditions were template denaturation and polymerase activation at 95°C for 2 min followed by 45 cycles of 95°C denaturation for 15 sec, 60°C annealing and extension for 45 sec with a single fluorescence measurement at each cycle. When the reaction was completed, melting curves were plotted to confirm the specificity of the product. Data were normalized based on the sets of reference genes (H2afz, Hprt1 and Ywhaz) [18]. Similar normalization procedures during mouse *Pou5f1* transcript quantification were carried out using sets of mouse reference genes are published previously [47].

Statistical analysis

Data were analyzed using the Chi-square test. P-values less than 0.05 were considered statistically significant.

Authors' contributions

JK jointly conceived the studies and performed the experimental design, carried out subcloning of promoter regions, performed the functional promoter test in mouse ESCs, and wrote the manuscript. KK carried out the molecular genetic studies and participated in the sequence alignment. ZP prepared *in vivo* and *in vitro* rabbit embryos for the experiments. SM carried out the real-time RT-PCRs analyses, and participated in manuscript preparation. CRG carried out the rabbit BAC library screens and provided the clones. ZT participated in cDNA cloning from blastocyst stage rabbit embryos. IB participated in computational analysis of DNA sequences, pseudogene specific primer design, and qPCRs. ABG participated in primer design, data analyses, and performed statistical tests. KT contributed to the experimental manipulations, MKP gave constructive comments on the study design and drafting the manuscript, AD conceived the studies, supervised the study design, execution, analysis, and approved the final version. All authors read and approved the manuscript.

Additional material

Additional file 1

Pseudogene of rabbit POU5F1. The data provided represent the alignment of rabbit cDNA and pseudogene sequence, and the expression of the pseudogene in preimplantation stage rabbit embryos.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2199-10-88-S1.pdf>]

Acknowledgements

The authors would like to thank Mrs Hajnalka Csákány Tolnainé for her excellent technical assistance during the experiment and for Dr. Claudia Stanca for her advice and help with the preimplantation-stage rabbit embryo manipulations and expression studies. This research was supported by Wellcome Trust (Grant No. 070246); OTKA T046171, OMFB-01220/2006, NKTH-OTKA FP7 Mobility "HUMAN-MB08C-80205", Hungarian-Chinese NKTH TET (No. CN-56/2007) projects; NKTH/ANR TET Franco-Hungarian Bilateral Scientific and Technological Collaborative Project "Plurabit", NKTH/KPI Kozma F. TUDAS-I-2006-0005, and European Union FP6 CLONET (MRTN-CT-2006-035468) and MED-RAT (LSHG-CT-2006-518240).

References

- Scholer HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P: **New type of POU domain in germ line-specific protein Oct-4.** *Nature* 1990, **344**:435-9.
- Yeom YI, Fuhrmann G, Ovitt CE, Brehm A, Ohbo K, Gross M, Hubner K, Scholer HR: **Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells.** *Development* 1996, **122**:881-94.

3. Kirchhof N, Carnwath JW, Lemme E, Anastasiadis K, Scholer H, Niemann H: **Expression pattern of Oct-4 in preimplantation embryos of different species.** *Biol Reprod* 2000, **63**:1698-705.
4. van Eijk MJ, van Rooijen MA, Modina S, Scesi L, Folkers G, van Tol HT, Bevers MM, Fisher SR, Lewin HA, Rakacolli D, et al.: **Molecular cloning, genetic mapping, and developmental expression of bovine POU5F1.** *Biol Reprod* 1999, **60**:1093-103.
5. Cauffman G, Velde H Van de, Liebaers I, Van Steirteghem A: **Oct-4 mRNA and protein expression during human preimplantation development.** *Mol Hum Reprod* 2005, **11**:173-81.
6. Mitalipov SM, Kuo HC, Hennebold JD, Wolf DP: **Oct-4 expression in pluripotent cells of the rhesus monkey.** *Biol Reprod* 2003, **69**:1785-92.
7. Boiani M, Scholer HR: **Regulatory networks in embryo-derived pluripotent stem cells.** *Nat Rev Mol Cell Biol* 2005, **6**:872-84.
8. Cauffman G, Liebaers I, Van Steirteghem A, Velde H Van de: **POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos.** *Stem Cells* 2006, **24**:2685-91.
9. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al.: **Core transcriptional regulatory circuitry in human embryonic stem cells.** *Cell* 2005, **122**:947-56.
10. Pesce M, Anastasiadis K, Scholer HR: **Oct-4: lessons of totipotency from embryonic stem cells.** *Cells Tissues Organs* 1999, **165**:144-52.
11. Nordhoff V, Hubner K, Bauer A, Orlova I, Malapetsa A, Scholer HR: **Comparative analysis of human, bovine, and murine Oct-4 upstream promoter sequences.** *Mamm Genome* 2001, **12**:309-17.
12. Ovitt CE, Scholer HR: **The molecular biology of Oct-4 in the early mouse embryo.** *Mol Hum Reprod* 1998, **4**:1021-31.
13. Fang ZF, Gai H, Huang YZ, Li SG, Chen XJ, Shi JJ, Wu L, Liu A, Xu P, Sheng HZ: **Rabbit embryonic stem cell lines derived from fertilized, parthenogenetic or somatic cell nuclear transfer embryos.** *Exp Cell Res* 2006, **312**:3669-82.
14. Wang S, Tang X, Niu Y, Chen H, Li B, Li T, Zhang X, Hu Z, Zhou Q, Ji W: **Generation and characterization of rabbit embryonic stem cells.** *Stem Cells* 2007, **25**:481-9.
15. Meng Q, Polgar Z, Liu J, Dinnyes A: **Live birth of somatic cell-cloned rabbits following trichostatin A treatment and cotransfer of parthenogenetic embryos.** *Cloning Stem Cells* 2009, **11**:203-208.
16. Chesne P, Adenot PG, Viglietta C, Baratte M, Boulanger L, Renard JP: **Cloned rabbits produced by nuclear transfer from adult somatic cells.** *Nat Biotechnol* 2002, **20**:366-9.
17. Shi JJ, Cai DH, Chen XJ, Sheng HZ: **Cloning and characterization of the rabbit POU5F1 gene.** *DNA Seq* 2008, **19**:56-61.
18. Mamo S, Gal AB, Polgar Z, Dinnyes A: **Expression profiles of the pluripotency marker gene POU5F1 and validation of reference genes in rabbit oocytes and preimplantation stage embryos.** *BMC Mol Biol* 2008, **9**:67.
19. Rogel-Gaillard C, Piumi F, Billault A, Bourgeaux N, Save JC, Urien C, Salmon J, Chardon P: **Construction of a rabbit bacterial artificial chromosome (BAC) library: application to the mapping of the major histocompatibility complex to position 12q.I.1.** *Mamm Genome* 2001, **12**:253-5.
20. Yeom YI, Ha HS, Balling R, Scholer HR, Artzt K: **Structure, expression and chromosomal location of the Oct-4 gene.** *Mech Dev* 1991, **35**:171-9.
21. Hubner K, Fuhrmann G, Christenson LK, Kehler J, Reinbold R, De La Fuente R, Wood J, Strauss JF 3rd, Boiani M, Scholer HR: **Derivation of oocytes from mouse embryonic stem cells.** *Science* 2003, **300**:1251-6.
22. Brunet-Simon A, Henrion G, Renard JP, Duranthon V: **Onset of zygotic transcription and maternal transcript legacy in the rabbit embryo.** *Mol Reprod Dev* 2001, **58**:127-36.
23. Leandri RD, Archilla C, Bui LC, Peynot N, Liu Z, Cabau C, Chastellier A, Renard JP, Duranthon V: **Revealing the dynamics of gene expression during embryonic genome activation and first differentiation in the rabbit embryo with a dedicated array screening.** *Physiol Genomics* 2009, **36**:98-113.
24. Mighell AJ, Smith NR, Robinson PA, Markham AF: **Vertebrate pseudogenes.** *FEBS Lett* 2000, **468**:109-14.
25. Vanin EF: **Processed pseudogenes: characteristics and evolution.** *Annu Rev Genet* 1985, **19**:253-72.
26. Pugh BF, Tjian R: **Transcription from a TATA-less promoter requires a multisubunit TFIID complex.** *Genes Dev* 1991, **5**:1935-45.
27. Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, et al.: **Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells.** *Mol Cell Biol* 2005, **25**:6031-46.
28. Pan GJ, Chang ZY, Scholer HR, Pei D: **Stem cell pluripotency and transcription factor Oct4.** *Cell Res* 2002, **12**:321-9.
29. Kurosaka S, Eckardt S, McLaughlin KJ: **Pluripotent lineage definition in bovine embryos by Oct4 transcript localization.** *Biol Reprod* 2004, **71**:1578-82.
30. Bellingham J, Gregory-Evans K, Fox MF, Gregory-Evans CY: **Gene structure and tissue expression of human selenoprotein W, SEPWI, and identification of a retroprocessed pseudogene, SEPWI P.** *Biochim Biophys Acta* 2003, **1627**:140-6.
31. Hirotsune S, Yoshida N, Chen A, Garrett L, Sugiyama F, Takahashi S, Yagami K, Wynshaw-Boris A, Yoshiki A: **An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene.** *Nature* 2003, **423**:91-6.
32. Korneev SA, Park JH, O'Shea M: **Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene.** *J Neurosci* 1999, **19**:7711-20.
33. Elliman SJ, Wu I, Kemp DM: **Adult tissue-specific expression of a Dppa3-derived retrogene represents a postnatal transcript of pluripotent cell origin.** *J Biol Chem* 2006, **281**:16-9.
34. Pain D, Chirn GW, Strassel C, Kemp DM: **Multiple retroseudogenes from pluripotent cell-specific gene expression indicates a potential signature for novel gene identification.** *J Biol Chem* 2005, **280**:6265-8.
35. Takeda J, Seino S, Bell GI: **Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues.** *Nucleic Acids Res* 1992, **20**:4613-20.
36. Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H: **A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.** *Cell* 1990, **60**:461-72.
37. Okazawa H, Okamoto K, Ishino F, Ishino-Kaneko T, Takeda S, Toyoda Y, Muramatsu M, Hamada H: **The oct3 gene, a gene for an embryonic transcription factor, is controlled by a retinoic acid repressible enhancer.** *Embo J* 1991, **10**:2997-3005.
38. Panagopoulos I, Moller E, Collin A, Mertens F: **The POU5F1P1 pseudogene encodes a putative protein similar to POU5F1 isoform I.** *Oncol Rep* 2008, **20**:1029-33.
39. Liedtke S, Enczmann J, Waclawczyk S, Wernet P, Kogler G: **Oct4 and its pseudogenes confuse stem cell research.** *Cell Stem Cell* 2007, **1**:364-6.
40. Liedtke S, Stephan M, Kogler G: **Oct4 expression revisited: potential pitfalls for data misinterpretation in stem cell research.** *Biological Chemistry* 2008, **389**:845-850.
41. Monk M, Hitchins M, Hawes S: **Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development.** *Mol Hum Reprod* 2008, **14**:347-55.
42. **GADIE Biologicals Resources Centre** [<http://www-crb.jouy.inra.fr/BRC/libraries-e.htm>]
43. **ClustalW2 program** [<http://www.ebi.ac.uk/Tools/clustalw2/>]
44. **EMBOSS Pairwise Alignment Algorithms program** [<http://www.ebi.ac.uk/Tools/emboss/align/>]
45. Mitalipov SM, White KL, Farrar VR, Morrey J, Reed WA: **Development of nuclear transfer and parthenogenetic rabbit embryos activated with inositol 1,4,5-trisphosphate.** *Biol Reprod* 1999, **60**:821-7.
46. Tanaka N, Takeuchi T, Neri QV, Sills ES, Palermo GD: **Laser-assisted blastocyst dissection and subsequent cultivation of embryonic stem cells in a serum/cell free culture system: applications and preliminary results in a murine model.** *J Transl Med* 2006, **4**:20.
47. Mamo S, Gal AB, Bodo S, Dinnyes A: **Quantitative evaluation and selection of reference genes in mouse oocytes and embryos cultured in vivo and in vitro.** *BMC Dev Biol* 2007, **7**:14.
48. Nagy A, Rossant J, Nagy R, Abramow-Newerly W, Roder JC: **Derivation of completely cell culture-derived mice from early-**

- passage embryonic stem cells.** *Proc Natl Acad Sci USA* 1993, **90**:8424-8.
49. Nagy A: **Manipulating the mouse embryo: a laboratory manual.** 3rd edition. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2003.
50. **BLASTn program** [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>]
51. Gal AB, Carnwath JW, Dinnyes A, Herrmann D, Niemann H, Wrenzycki C: **Comparison of real-time polymerase chain reaction and end-point polymerase chain reaction for the analysis of gene expression in preimplantation embryos.** *Reprod Fertil Dev* 2006, **18**:365-71.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

