STEM CELLS®
Original Article

This material is protected by U.S. Copyright law.
Unauthorized reproduction is prohibited.
For reprints contact: Reprints@AlphaMedPress.com

Human *STELLAR*, *NANOG*, and *GDF3* Genes Are Expressed in Pluripotent Cells and Map to Chromosome 12p13, a Hotspot for Teratocarcinoma

AMANDER T. CLARK,[a-d] RYAN T. RODRIGUEZ,[a] MEGAN S. BODNAR,[a] MICHAEL J. ABEYTA,[a-d] MARCELLE I. CEDARS,[a] PAUL J. TUREK,[a,c] MERI T. FIRPO,[a] RENEE A. REIJO PERA[a-d]

[a]Department of Obstetrics, Gynecology, and Reproductive Sciences, Center for Reproductive Sciences, San Francisco, California; [b]Departments of Physiology and [c]Urology, [d]Programs in Human Genetics, Cancer Genetics and Development, and Stem Cell Biology, University of California at San Francisco, San Francisco, California, USA

Key Words. *Human embryonic stem cells · Pluripotent · Germ cell* · OCT-4 · STELLAR · NANOG · GDF3 · *Ovary · Testis · Embryonal carcinoma · Teratoma*

ABSTRACT

Genes required to maintain pluripotency in human embryonic stem (hES) cells are largely unknown, with the exception of *OCT-4*, a homolog of mouse *Oct-4*, which is critical for the establishment of the embryonic inner cell mass and the generation of totipotent mouse embryonic stem (mES) cell lines. In the current study, we identified two genes with expression similar to *OCT-4*, in that they are largely restricted to pluripotent hES cells, premeiotic germ lineage cells, and testicular germ cell tumor cells. Furthermore, we determined that upon hES cell differentiation, their expression is downregulated. The genes we identified in the current study include the human *stella*-related (*STELLAR*) gene, which encodes a highly divergent protein (with just 32.1% identity to mouse *stella* over the 159 amino acid sequence) that maps to human chromosome 12p13. Notably, human *STELLAR* is located distal to a previously uncharacterized homeobox gene, which is the human homolog of the recently identified murine gene, *Nanog*, and proximal to the *GDF3* locus, whose transcription is restricted to germ cell tumor cells. Our characterization of *STELLAR*, *NANOG*, and *GDF3* suggests that they may play a similar role in humans as in mice, in spite of their remarkable evolutionary divergence. Stem Cells *2004;22:169-179*

INTRODUCTION

Little is known of the determinants of pluripotency in human cells. However, one of the best characterized genes is the POU-domain class-5 transcription factor, *POU5f1* (*OCT-4*). *OCT-4* is a classic marker of pluripotency in human embryonic stem (hES) cells; expression has been shown to coordinately decrease with differentiation in vitro into embryoid bodies (EBs) and in vivo with the formation of teratocarcinomas [1]. Additional pluripotency factors have not yet been identified in humans, though several

Correspondence: Renee A. Reijo Pera, Ph.D., Center for Reproductive Sciences, HSE1656, Box 0556 UCSF, San Francisco, California 94143-0556, USA. Telephone: 415-476-3178; Fax: 415-476-3121; e-mail: reijo@itsa.ucsf.edu *Received July 22, 2003; accepted for publication October 20, 2003.* *©AlphaMed Press 1066-5099/2004/$12.00/0*

candidate factors that function in murine ES (mES) cells, such as *Fgf4*, *Foxd3*, and the recently identified *Nanog* genes, have been reported [2-5]. Although we might assume that a program similar to that in mES cells is required for maintenance of hES cell pluripotency and self-renewal, there are some obvious differences between the two mammalian systems. For example, it is well known that the critical leukemia inhibitory factor-signaling pathways are required for undifferentiated self-renewal of mES cells under normal physiological conditions [6], yet they are not sufficient for the same function in hES cells [1, 7, 8]. As a result, novel approaches will be necessary in order to identify the candidate genes associated with the maintenance of hES cell pluripotency.

A common biological link among hES cells, the epiblast, and premeiotic germ cells is pluripotency, or the ability of these cell types to contribute to multiple embryonic lineages [1, 9]. Interestingly, *Oct-4* is enriched in each of these three cell populations in mice; with respect to germ cells, expression is reduced as meiotic differentiation is initiated [10]. Thus, we hypothesized that genes with restricted expression patterns similar to *OCT-4* also may be critical regulators of pluripotency in humans.

Examination of Unigene EST databases for tissues that express human *OCT-4* revealed that it is largely expressed in human germ cell tumors. The elevated expression may reflect an increase in the germ cell (*OCT-4*$^+$) to somatic cell (*OCT-4*$^-$) ratio in these tumors. Based on these observations, we sought to identify novel genes that are expressed exclusively in pluripotent cell types by taking advantage of an observation that a consistent structural chromosomal abnormality associated with the formation and/or overproliferation of germ cell tumors in men is the formation of 12p isochromosomes. These isochromosomes may overexpress 12p genes that are associated with excessive growth of undifferentiated germ lineage cells [11, 12]. As a result, we chose to identify novel genes that may be associated with the molecular regulation of pluripotency by focusing on human chromosome 12p. The region of the mouse genome syntenic to human 12p is the distal end of mouse chromosome 6.

A search for loci present on the mouse chromosome 6 identified *stella* [13]. *stella* encodes a putative DNA-binding protein that is expressed differentially between nascent germ cells and their somatic neighbors at E7.25 in mice. As development progresses, *stella* expression is restricted to germ line stem cells. Given that mouse *stella* passed our criteria of having Oct-4-like expression and mapped to a chromosomal location syntenic to human chromosome 12p, we explored this genomic locus in more detail.

Materials and Methods

Bioinformatics

The genomic structure of mouse *stella* was obtained by Basic Local Alignment Sequence Tool (BLAST) analysis of the National Center for Biotechnology Information (NCBI) High Throughput Genomic Sequences (HTGS) database with the *stella* cDNA sequence (AY082485) and confirmed by analysis of the homologous mouse BAC (RP23117I23) using ENSEMBL (http://www.ensembl.org). The mouse *stella* locus, 2410075G02Rik (NM_139218), is also known as *PGC7* (AB072734) and *Dppa3* (NM_139218, G73534, AF490347). The syntenic human *stella* locus was identified by BLAST analysis of mouse *stella* in the human ENSEMBL database.

We identified a predicted gene with homology to mouse *stella* (ENS0000034235) on human chromosome 12p. Full-length human STELLA cDNA was obtained by reverse-transcription polymerase chain reaction (RT-PCR) from testis and ovary cDNA. Genomic structure of the predicted STELLA-related (*STELLAR*) gene was determined by BLAST analysis in ENSEMBL. Identification of the 3′ end of *STELLAR* was obtained by 3′ rapid amplification of cDNA ends (RACE) from testis cDNA using the SMART/CDSIII cDNA library construction kit (Clontech Inc.; Palo Alto, CA; http://www.clontech.com/index.shtml). BLAST analysis of the human *STELLAR* nucleotide sequence in the NCBI database also identified predicted genomic pseudogenes (intronless genes) on human chromosome 12, 14, and X; there were no corresponding mouse pseudogenes identified to any of these syntenic locations in the mouse genome.

Cell Culture

Information regarding the hES cell lines HSF-6, HSF-1, and H9 can be obtained at http://stemcells.nih.gov/stemcell. The ES lines, HSF-6 and HSF-1, were obtained from the University of California, San Francisco (UCSF; National Institutes of Health [NIH] code UCO6 and UC01, respectively). H9 (NIH code W-9) was obtained from the University of Wisconsin [1]. Undifferentiated hES cell colonies were cultured on irradiated CF1 mouse embryonic fibroblast feeder cells at 5% CO_2 in medium containing knock-out Dulbecco's-modified Eagle's medium (DMEM)-high glucose, 20% KnockOut Serum Replacer, 1 mM glutamine, nonessential amino acids (all from GIBCO BRL; Carlsbad, CA; http://www.invitrogen.com), 0.1 mM β-mercaptoethanol (Sigma; St. Louis, MO; http://www.sigmaaldrich.com), and 4 ng/μl fibroblast growth factor-2 (R&D Systems; Minneapolis, MN; http://www.rndsystems.com). To differentiate hES cells into embryoid bodies (EBs), cells were cultured at 5% CO_2 in medium containing knockout DMEM-high glucose, 20% fetal calf serum, 1 mM glutamine,

and 0.1 mM β-mercaptoethanol. At days 0, 3, 7, and 14, EBs were collected, centrifuged, and resuspended in 600 ml RLT buffer (Qiagen; Valencia, CA; http://www.qiagen.com).

RNA and cDNA Production

Adult human testis mRNA was from Clontech Inc. Adult female ovary mRNA was from Ambion (Austin, TX; http://www.ambion.com). cDNA was generated from mRNA using 250 ng of random hexamers under standard conditions with murine leukemia virus RT (Promega; Madison, WI; http://www.promega.com). Seminoma samples were obtained from the UCSF Cancer Center Tissue Registry; testicular biopsies and oocytes were from the UCSF Center for Reproductive Health.

All human samples were obtained after Institutional Review Board approval. Total RNA isolated from all tissues and cells, with the exception of oocytes, was extracted via the RNeasy system according to instructions (Qiagen). cDNA was generated via random priming as above. Total RNA from oocytes was extracted using the PicoPure RNA isolation system (Arcturus; Mountain View, CA; http://www.arctur.com) followed by cDNA production with 250 ng/μl random hexamers (Promega) as above. Following extraction from human oocytes, cDNA was concentrated using DNA Clean and Concentrator (Zymo Research; Orange, CA; http://www.zymoresearch.com). cDNA from pooled human fetal ovaries at 20-29 weeks of gestation was obtained from Spring Bioscience (Fremont, AZ; http://www.springbio.com), and the fetal cDNA panel was from Clontech. PCR was performed with 50 ng or 100 ng of first-strand cDNA reaction as specified.

Polymerase Chain Reaction

Polymerase chain reactions contained 3 mM $MgCl_2$; 10 mM each of dATP, dGTP, dCTP, and dTTP; 2 μM primers; and 0.25 U platinum Taq (Invitrogen; Carlsbad, CA; http://www.invitrogen.com). The PCR reaction was initiated by hot start at 95ºC for 5 minutes followed by 35 cycles at 95ºC for 30 seconds, 60ºC for 30 seconds, and 72ºC for 30 seconds. Real-time PCR for human *STELLAR*, *GDF3*, *NANOG*, *SYCP1*, and *OCT-4* was performed as above except for addition of 1× SYBR Green (Molecular Probes; Eugene, OR; http://www.probes.com), 1× fluorescein (Bio-Rad; Hercules, CA; http://www.bio-rad.com), and 2% dimethylsulfoxide. All reactions for real-time PCR were performed in the presence of 4.5 mM $MgCl_2$ and analyzed using an iCycler iQ (Bio-Rad), calibrated for use with SYBR green. Equal reaction efficiencies of all five genes were verified, and mean normalized expression was calculated using Relative Expression Software Tool (REST XL) [14].

All experiments included negative controls with no cDNA, and primers were designed to span exons to distinguish cDNA. Primers were: *STELLAR* primers: (5′-GTTACTGGGCGGAGTTCGTA-3′)/(5′-TGAAGTGGCTTGGTGTCTTG-3′) (174 bp); *GDF3* primers: (5′-AGACTTATGCTACGTAAAGGAGCT-3′)/(5′-CTTTGATGGCAGACAGGTTAAAGTA-3′) (150 bp); *NANOG* primers: (5′-CAGCTGTGTGTACTCAATGATAGATTT-3′)/(5′-CAACTGGCCGAAGAATAGCAATGGTGT-3′) (142 bp); *Deleted in AZoospermia-Like* (*DAZL*) primers: (5′-ATGTTAGGATGGATGAAACTGAGATTA-3′)/(5′-CCATGGAAATTTATCTGTGATTCTACT-3′) (178 bp); GAPDH primers: (5′-ACCACAGTCCATGCCATCAC-3′)/(5′-TCCACCACCCTGTTGCTGTA-3′) (500 bp); *OCT-4* primers: (5′-ACATCAAAGCTCTGCAGAAAGAACT-3′)/(5′-CTGAATACCTTCCCAAATAGAACCC-3′) (133 bp); NCAM-1 primers: (5′-ATGGAAACTCTATTAAAGTGAACCTGA-3′)/(5′-TAGACCTCATACTCAGCATTCCAGT-3′) (187 bp); and *SYCP1* primers: (5′-AAGATTTACAGATAGCAACAAACACA-3′)/(5′-AATCTTTGCTGTTCTGTTCTCAATAA-3′) (169 bp).

Northern Blot

Purified *STELLAR* (nucleotides 362-444; AY230136) and *NANOG* (nucleotides, 115-413; AY230262) PCR products were labeled via random priming per the manufacturer's instructions (Roche; Indianapolis, IN; http://www.roche.com) and incorporation of freshly labeled α^{32}P dCTP. Human Northern membranes and hybridization solution were purchased from Clontech for analysis of *STELLAR* and from RNWAY Laboratories Inc. (Seoul, Korea; http://www.rnway.com) for analysis of *NANOG*. Membranes were hybridized and washed as previously outlined [15] and exposed to XOMAT film at -80ºC for 10 days. The membranes were stripped then reprobed with labeled GAPDH probe to ensure equal loading of samples.

Statistics

Pair-wise comparisons were made using analysis of variance, followed by two-tailed *t*-tests. Significance was assumed at $p < 0.05$.

Results

Characterization of Syntenic Human *STELLAR* and Surrounding Genomic Sequences

We localized the mouse genomic locus of *stella* to BAC RP23-180N22 (house mouse) and RP23-117I23 (C57BL/6J) by BLAST analysis of the NCBI HTGS database. All predicted and known sequences homologous with BAC RP23-117I23 and RP23-180N22 were then identified using ENSEMBL. From this analysis, a RIKEN deposited sequence (2410075G02Rik [NM_139218]) corresponding

to the gene *PGC7* [16] was also identified. Using synteny maps in the Mouse Genome Informatics and Mouse Genome Resource databases, together with the location of the neighboring MIT marker D6MIT151 as an anchor, we identified the region of the short arm of human chromosome 12p, which was syntenic to the mouse *stella* locus (mouse chromosome 6 at 59.3cM). Analysis of the human map at this site (BAC RP11-277J24) revealed that the homologous region contained a predicted gene that had no previous experimental evidence for transcription. Analysis of this gene revealed that, like mouse *stella* (Fig. 1A), the predicted syntenic human gene (which we called *STELLAR*) contains four exons of length similar to the mouse gene (Fig. 1A). Further sequence analysis of BAC RP11-277J24 revealed a novel locus 72.5 kb distal to *STELLAR* called *FLJ12581 fis* (AK022643; Unigene Hs.329296) (Fig. 1B).

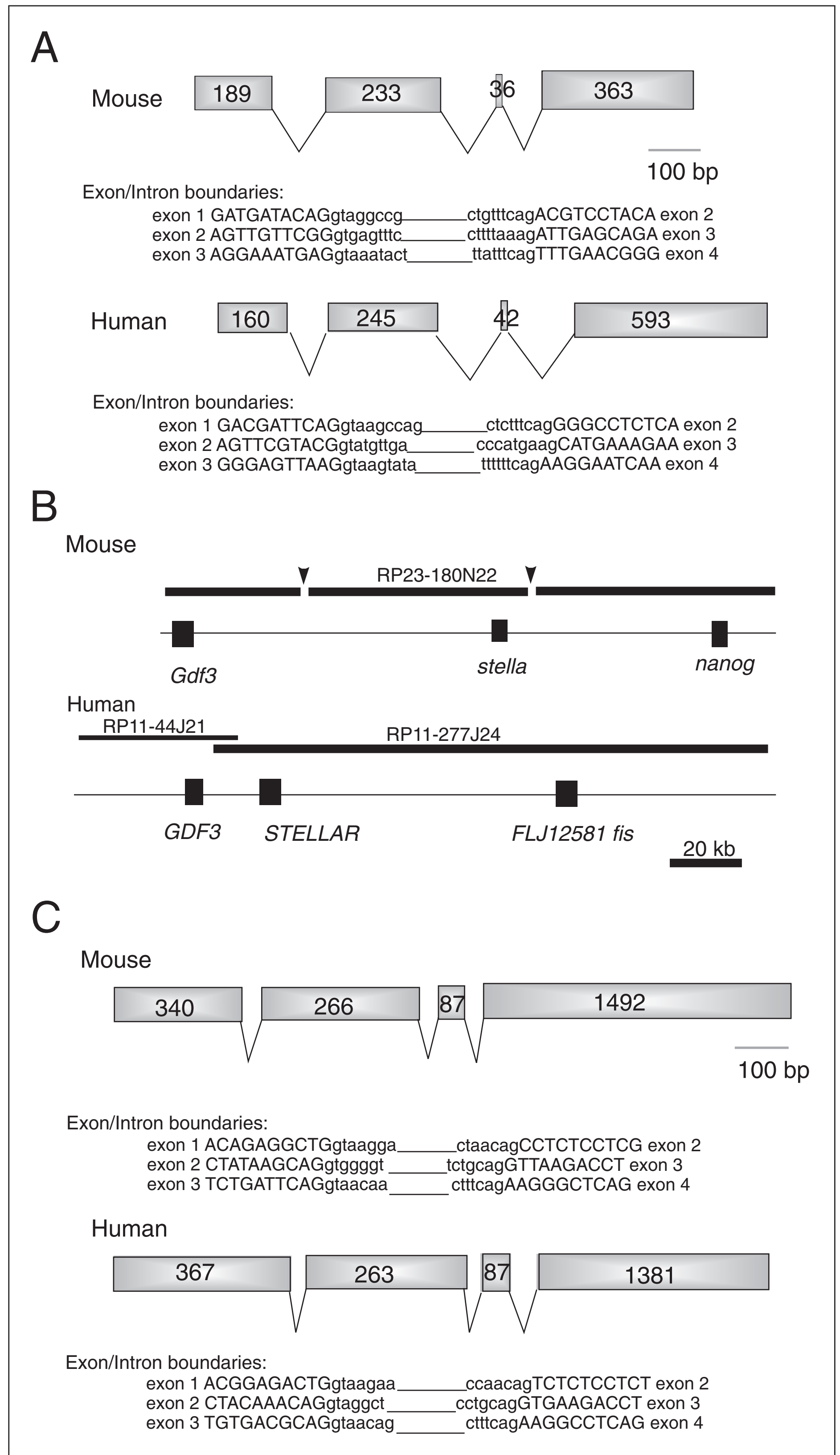

Figure 1. Genomic organization and exon/intron structure of the mouse stella and human* STELLAR *locus on mouse chromosome 6 and human chromosome 12, respectively (A). *Both mouse* stella *and human* STELLAR *are composed of four exons and the exon/intron boundaries are shown. B) The mouse physical map in ENSEMBL reveals that the* stella *locus is approximately 71.4 kb distal to the mouse* Gdf3 *locus and 53.7 kb proximal to* nanog. *Mouse BAC RP23-117I23 contains two sequence gaps represented by arrow heads. In humans, the* STELLAR *locus is 15.6 kb distal to* GDF3 *and 72.5 kb proximal to the predicted gene,* FLJ12581 fis *(*NANOG*), which is syntenic to the* RIKEN *sequence 241000E02Rik (*nanog*). The two complementing human BACs (RP11-44J21 and RP11-277J24), which were used to generate the human genomic contig, are fully sequenced. C) Genomic organization and exon/intron structure of the mouse* nanog *and human* NANOG *locis. Both syntenic loci are composed of four exons. Nucleotide sequences of exon/intron junctions are shown. Uppercase letters refer to exon sequences. Intron sequences are represented by lowercase letters. All introns begin with the dinucleotide "gt" and end with the dinucleotide "ag" at both the mouse and human locis. (______) represents intronic sequence between each boundary.*

We originally named this novel human locus Neighbor Of *STELLAR* (NOSTEL); however, recent work has identified this novel locus as the human homolog of the mouse *nanog* gene [5]. The proximal end of BAC RP11-277J24 hybridizes with a second BAC called RP11-44J21, which contains the growth and differentiation factor 3 (*Gdf3*) locus (Fig. 1B). By forming a contig between the two BACs, it can be determined that human *STELLAR* is 15.6 kb distal to *GDF3* (Fig. 1B). In the mouse, *Gdf3*, *stella*, and *nanog* are represented in BAC RP23-117I23. The current ENSEMBL order is shown in Figure 1B; however, as more sequence information becomes available, the distances between *STELLAR*, *NANOG*, and *GDF3* in the mouse genome will potentially change. Currently, *stella* is placed 71.4 kb distal to *Gdf3* and 53.7 kb proximal to *nanog*. In the human genome, *NANOG* is annotated as a four-exon gene comparable to the mouse chromosome 6 homolog annotated in the ENSEMBL database (Fig. 1C).

To obtain the full-length cDNA sequence of human *STELLAR* and *NANOG*, we performed PCR amplification of human testis and hES cell (HSF-6) cDNA and compared the sequences of the cloned fragments with their respective annotated genomic loci. The accession numbers of cloned *STELLAR* and *NANOG* obtained in the current analysis are AY230136 and AY230262, respectively.

STELLAR and *NANOG* Expression in Adult and Embryonic Tissues

Northern blot analysis was used to examine expression of *NANOG* and *STELLAR* in adult human tissues (Fig. 2A and Fig. 2B). Expression of *GDF3* in adult human tissues was previously reported [17]. Expression of both *STELLAR* and *NANOG* in all human adult tissues examined was extremely low (Fig. 2A and 2B). In both cases, a minimum of 10 days of exposure was required in order to observe the weak signals. *STELLAR* probe hybridized to a 1.1-kb band in ovary, testis, and thymus consistent with the length of the cDNA transcript identified from human testis (Fig. 2A). In addition, a larger 4.4-kb *STELLAR* band of unknown origin was also detected in testis and ovary; 3′ RACE failed to identify this *STELLAR* transcript, although it may represent an alternatively spliced variant or novel 5′ structure not identified in the current analysis. *NANOG* was detected as a faint 2.2-kb band exclusively in the testis (Fig. 2B). Given this extremely low expression level, we used a more sensitive RT-PCR approach to examine expression of these three loci, and *OCT-4*, in human fetal tissues (Fig. 2C). We found that *OCT-4*, *STELLAR*, and *NANOG* were expressed only in fetal ovary. All other fetal tissues had basal levels of transcription comparable to *OCT-4* and lacked detectable expression. By comparison, *GDF3* was enriched in fetal

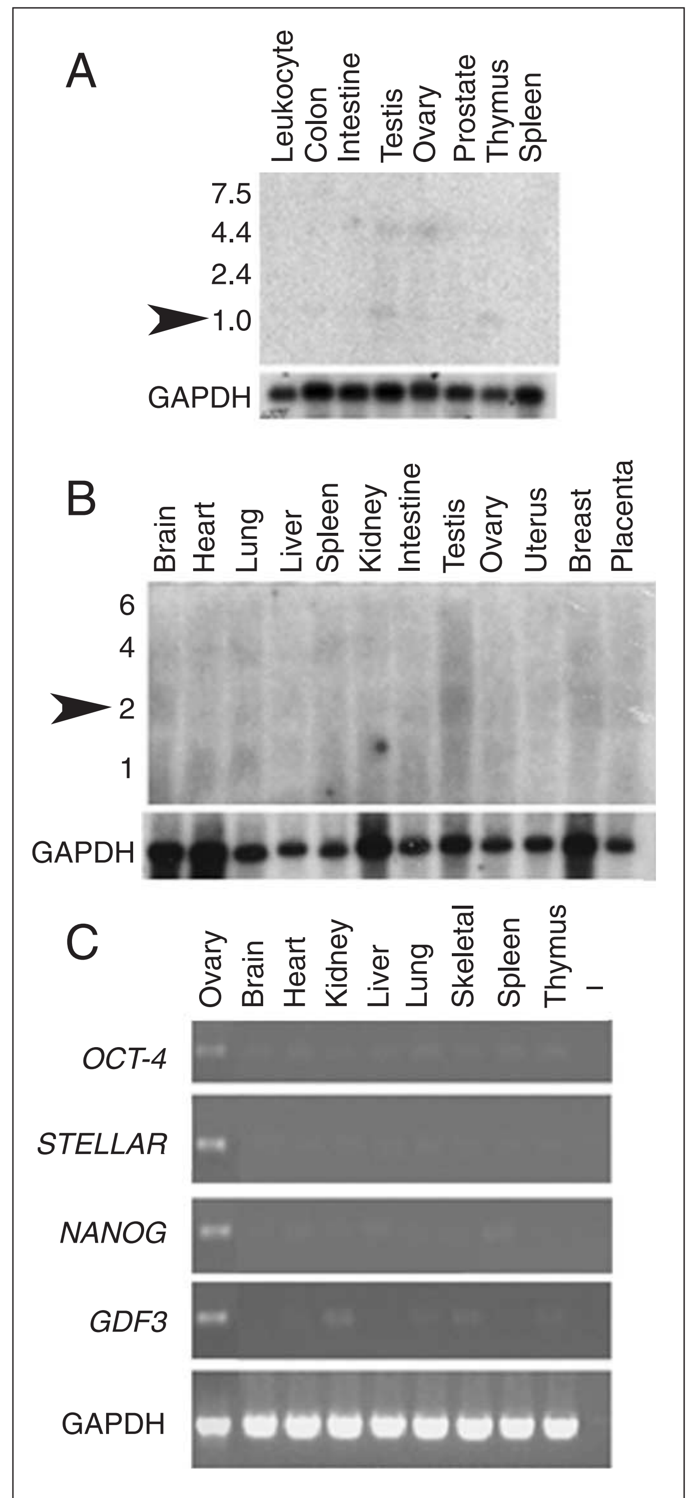

Figure 2. Northern blot analysis of normal human adult tissues and RT-PCR of human embryonic tissues. STELLAR *is expressed in testis, thymus, and ovary (A).* NANOG *is expressed exclusively in the adult human testis, albeit at very low levels (B). C) RT-PCR on mRNA isolated from human fetal tissues.* OCT-4, STELLAR, *and* NANOG *are expressed exclusively in the fetal ovaries at 20-29 weeks of gestation, whereas* GDF3 *is expressed in fetal ovary, kidney, skeletal muscle, and thymus.*

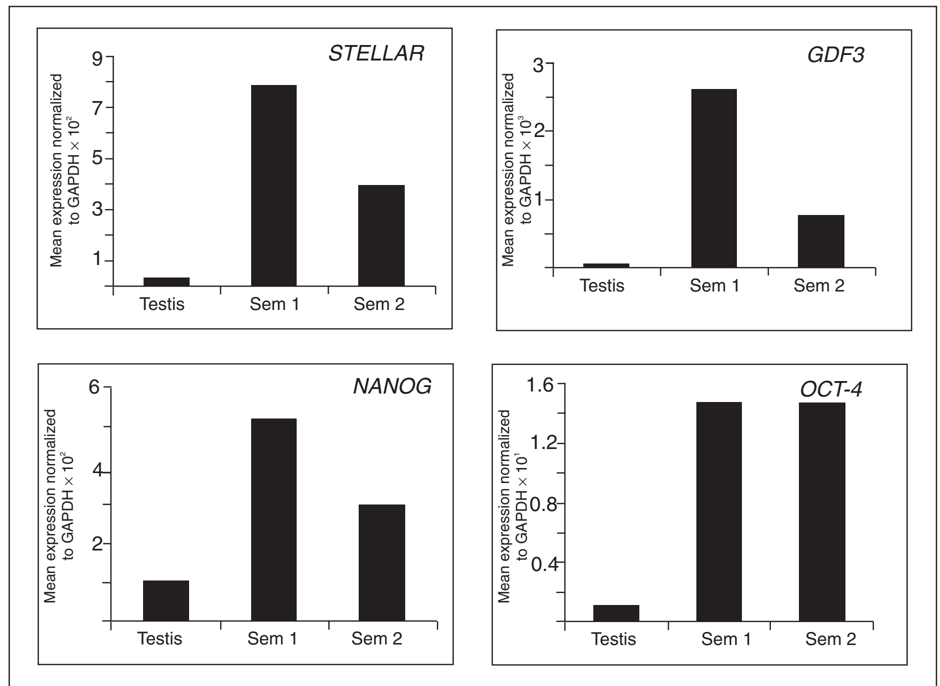

Figure 3. Real Time RT-PCR of* STELLAR *(A),* GDF3 *(B),* NANOG *(C), and* OCT-4 *(D) from two different seminoma specimens. *In all cases, the seminomas had greater expression than normal testis. 50 ng of first-strand cDNA were used in each reaction.*

ovaries; however, expression was also detected in fetal kidney, lung, skeletal muscle, and thymus.

Given that *GDF3* was previously cloned from human teratocarcinoma libraries [17] and human chromosome 12p is an apparent hotspot for chromosomal abnormalities associated with teratocarcinoma, we used real-time PCR analysis to compare expression of *STELLAR* and *NANOG* in normal human testis with that in two independently isolated testicular germ cell tumors (seminomas) (Fig. 3). We determined that transcription of all three genes was greater in germ cell tumors than in normal adult testis. In particular, *STELLAR* was greater by approximately seven- and fourfold, *GDF3* was greater by 12- and fourfold, and *NANOG* was greater by four- and threefold, respectively, in the two seminomas (Fig. 3). Furthermore, *OCT-4*, which in silico can be identified largely in germ cell tumor libraries, was greater by approximately ninefold compared with normal testis.

Amino Acid Identity of Human *STELLAR*

Translation of human *STELLAR* revealed a highly basic protein (pI = 8.3) of 159 amino acids, with a predicted mass of 17.9 kD (Fig. 4), similar to mouse *stella* (150 amino acids; 18 kD) [16]. Analysis of the human STELLAR peptide sequence using the pSORT prediction program suggests that human STELLAR, like mouse STELLA, has no known functional motifs. Furthermore, amino acid analysis predicted that, like mouse STELLA, the human syntenic STELLAR locus also encodes a nuclear protein. However, given these similarities, the degree of amino acid identity between the mouse and human proteins is remarkably low, with just 32.1% identity over the protein sequences. Comparisons of NANOG and GDF3 proteins with their respective mouse proteins have been described [5, 17] and also demonstrate low identity between these homologs (Table 1).

Human *STELLAR* and *NANOG* Are Enriched in Pluripotent Cell Populations

Given the expression of the *STELLAR*, *NANOG*, and *GDF3* genes, albeit at low levels in normal adult human testis and ovary, we sought to further characterize their expression in somatic cells and/or germ cells of the gonads. We compared expression of human *STELLAR*, *NANOG*, and *GDF3* in adult ovary, testis, and fetal ovary at 20-29

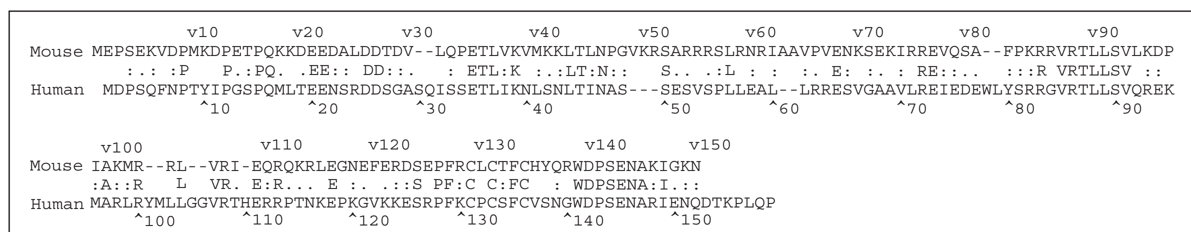

Figure 4. Amino acid sequence of mouse* stella *compared with human* STELLAR *cloned from a testis cDNA library with a predicted protein identity (mouse versus human) of 32.1% and similarity of 21.7%. *The symbols between each alignment indicate residue conservation, (letter for amino acid identity; colon for strongly conserved; and period for weakly conserved).*

Table 1. Germ cell and ES cell restricted loci at human chromosome 12p

Name (human)	Name (mouse)	Percentage homology (nt)	Percentage identity (peptide)	Category
GDF3	*Gdf3*	76.6%	69.3%	secreted
STELLAR	*Stella*	59.8%	32.1%	nuclear
NANOG	2410002E02Rik	59.9%	55.5%	nuclear

weeks of gestation. We also further defined expression in tissues with and without germ cells (Fig. 5A-5C). Our results show that all three genes and the positive control, *DAZL*, are expressed in the adult ovary; however, expression of *NANOG* in adult ovary was much lower than that of *STELLAR* or *GDF3* (Fig. 5A). Similarly, *GDF3* expression in the adult testis was lower than that of *STELLAR* and *NANOG*. RT-PCR on mRNA extracted from tissues of men who lack all germ cells (and were diagnosed with Sertoli cell only syndrome [SCOS]) was used to determine if *STELLAR*, *GDF3*, and *NANOG* are germ cell specific in the human testis (Fig. 5B).

Our results indicate that, in the absence of *DAZL* (in biopsies that lack germ cells), *STELLAR*, *GDF3*, and *NANOG* are also absent (Fig. 5B). Finally, we performed RT-PCR on ovulated, unfertilized, mature oocytes from women (Fig. 5C). We found that *STELLAR* was expressed in isolated oocytes together with the positive control, *DAZL* [18, 19], whereas, the somatic marker, neural cell adhesion molecule-1 (NCAM-1), was not expressed. Notably, unlike *STELLAR*, expression of both *GDF3* and *NANOG* were not detected in ovulated oocytes.

Human *STELLAR*, *NANOG*, and *GDF3* Expression Decreases in hES Cells with Differentiation

Given the similarities of expression of the *STELLAR*, *GDF3*, and *NANOG* genes to those of *OCT-4* in germ cells and germ cell tumors, we compared the relative levels of *STELLAR*, *GDF3*, and *NANOG* in three undifferentiated hES cell lines (HSF-6, HSF-1, and H9; Fig. 6). We found that all three genes were expressed in undifferentiated ES cells. *NANOG* was expressed at significantly higher levels than *STELLAR* and *GDF3* in HSF-1 (p = 0.023), H9 (p = 0.004), and HSF-6 (p = 0.002); there was no statistically significant difference in expression between *STELLAR* and *GDF3* in any of the three cell lines. In particular, *NANOG* was expressed at approximately 10-fold higher levels than *GDF3* and threefold higher levels than *STELLAR* in the hES cell line HSF-1 (Fig. 6B). Similar expression was observed in the H9 and HSF-6 cell lines.

In order to ensure that the hES cells in the current study were not expressing differentiated germ cell markers, we examined expression of the SYnaptonemal Complex 1

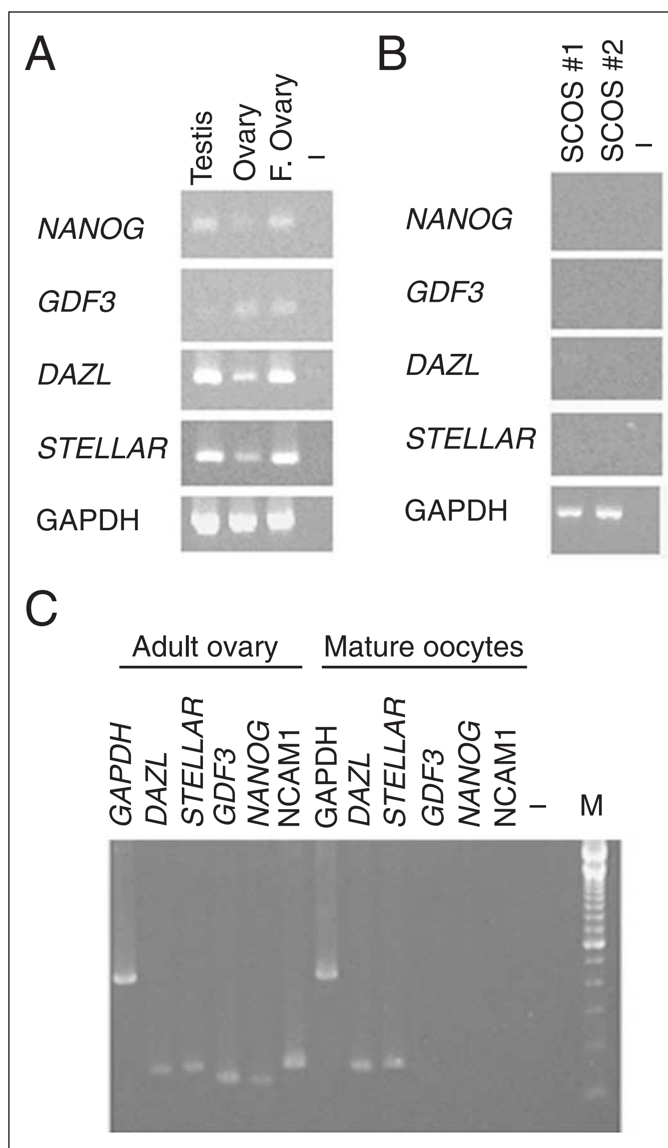

Figure 5. RT-PCR analysis showing expression of* NANOG, GDF3, *and* STELLAR *in human testis, ovary, and fetal ovary at 20-29 weeks of gestation (A). *Two samples from patients diagnosed with SCOS #1 and #2 were used to show that, in the absence of germ cells (therefore, absence of* DAZL*),* STELLAR, GDF3, *and* NANOG *are also absent (B). C) Expression in adult human ovary and isolated postovulatory oocytes.* DAZL, STELLAR, GDF3, NANOG, *and NCAM1 were all identified in the adult ovary, whereas only* STELLAR *and* DAZL *were expressed in mature oocytes (C). NCAM1 was used as a positive control for the presence of somatic tissue and a negative control for germ cells. GAPDH was used as a positive control for first strand cDNA production.*

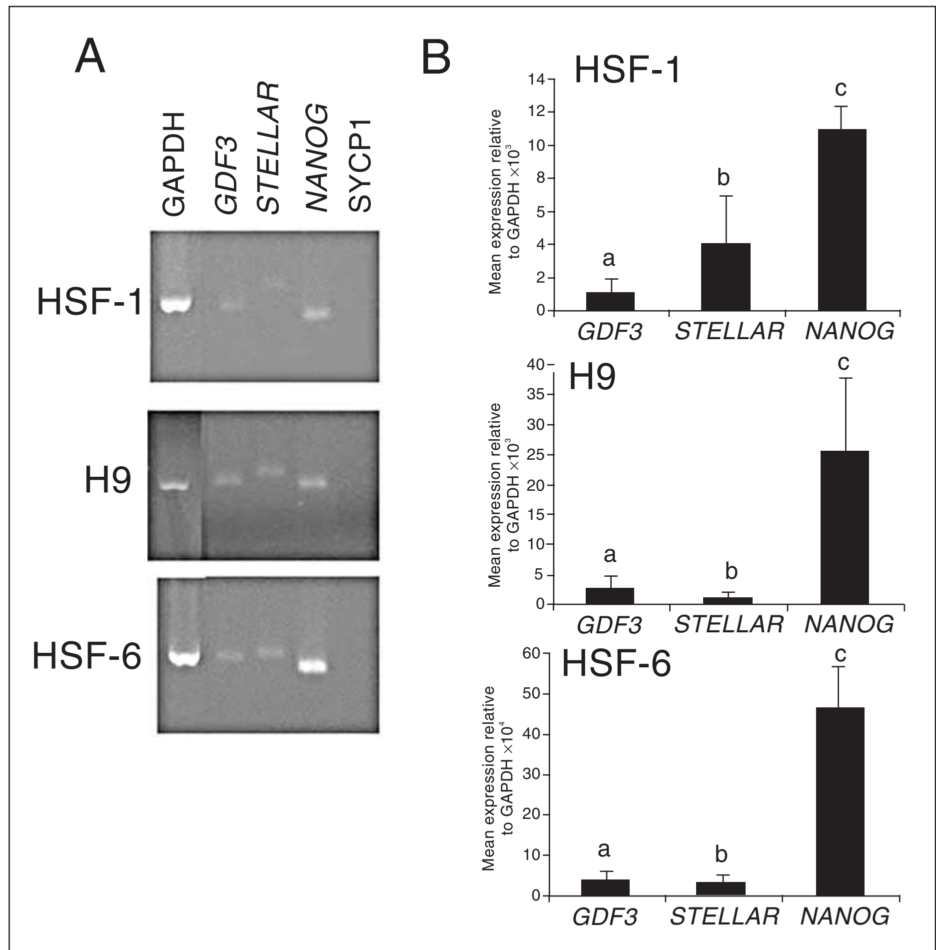

Figure 6. Real-time PCR was used to measure the relative levels of human* STELLAR, GDF3, *and* NANOG *in undifferentiated HSF-1, HSF-6, and H9 hES cells. *Mean normalized expression was calculated relative to GAPDH. Real-time PCR was performed in triplicate on three independently isolated samples. 50 ng of first-strand cDNA was used in each PCR reaction. The difference between A and B and A and C are significant.*

Protein (SYCP1), which is expressed only during meiosis I; it was not expressed at detectable levels (Fig. 6A). Finally, we examined expression of *STELLAR*, *GDF3*, and *NANOG* in differentiating hES cells in suspension culture at days 0, 3, 7, and 14 of EB formation (Fig. 7). Expression of *OCT-4*, *NANOG*, *GDF3*, and *STELLAR* were all decreased as differentiation progressed. In particular, *OCT-4* was significantly decreased by day 7 ($p = 0.003$), *NANOG* was significantly decreased by day 14 ($p = 0.02$), *STELLAR* was significantly decreased by day 3 ($p = 0.005$), and *GDF3* was significantly decreased by day 14 ($p = 0.009$). Although there was an initial apparent rise in *GDF3* expression on day 3 compared with day 0, this was not statistically significant. In contrast, the somatic lineage markers *NCAM-1* and α *fetoprotein* (*AFP*) were significantly ($p = 0.009$ and $p < 0.001$, respectively) increased with differentiation (Fig. 7).

Discussion

The restricted expression pattern of *Oct-4* during mouse embryo development, together with its critical role in maintaining mES cell self-renewal and pluripotency, prompted us to search for other loci in humans that share similar expression patterns to mouse *Oct-4*. Since an in silico search for *OCT-4* in human cDNA databases suggested that, in humans, transcripts with similar expression to *OCT-4* would be restricted to human germ cell tumors, we sought to identify additional loci with restricted human germ cell tumor expression.

We focused our search on novel loci on human chromosome 12p, which is a known chromosomal hotspot for structural chromosomal changes associated with formation of germ cell tumors. We identified the human *STELLAR* gene by synteny mapping to the mouse *stella* locus on mouse chromosome 6 and characterized the neighboring human genes, *NANOG* and *GDF3*, with regard to expression in adult and fetal tissues, as well as expression in pluripotent hES cells. We determined that human *STELLAR*, *NANOG*, and *GDF3* are enriched in organs containing pluripotent cells. In particular, we identified *NANOG* and *STELLAR* expression in fetal ovary with no detectable expression in any other fetal somatic tissue examined. By comparison, *GDF3* was enriched in fetal ovary, with lower levels of expression in fetal kidney, lung, skeletal muscle, and thymus. In adult human tissues, we determined that *STELLAR*, *NANOG*, and *GDF3* were localized to the ovary and testis, and expression of all three genes was upregulated in testicular germ cell tumor samples. In ovulated unfertilized oocytes, we found comparable levels of *STELLAR* and the germ cell-specific gene, *DAZL*, and no detectable expression of *NANOG* or *GDF3*. Analysis of three independently derived lines of undifferentiated pluripotent hES cells revealed that *STELLAR*, *GDF3*, and *NANOG* were consistently expressed in all three undifferentiated lines, and that expression decreased with differentiation.

This study using human tissues illustrates both the similarities and differences in gene expression of *GDF3*,

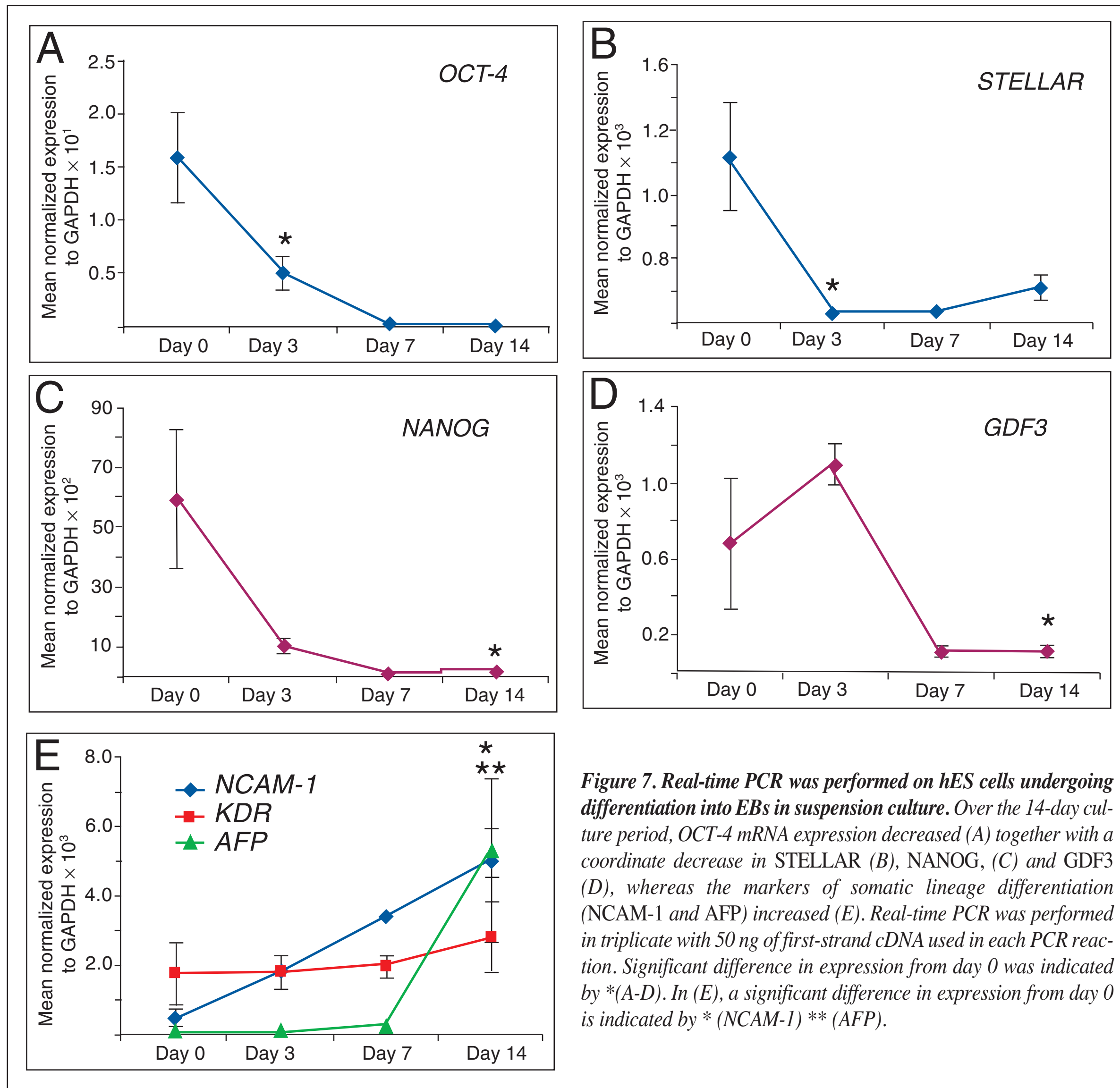

Figure 7. Real-time PCR was performed on hES cells undergoing differentiation into EBs in suspension culture. *Over the 14-day culture period, OCT-4 mRNA expression decreased (A) together with a coordinate decrease in* STELLAR *(B),* NANOG, *(C) and* GDF3 *(D), whereas the markers of somatic lineage differentiation (*NCAM-1 *and* AFP*) increased (E). Real-time PCR was performed in triplicate with 50 ng of first-strand cDNA used in each PCR reaction. Significant difference in expression from day 0 was indicated by *(A-D). In (E), a significant difference in expression from day 0 is indicated by * (NCAM-1) ** (AFP).*

STELLAR, and *NANOG* compared with previous reports from mice. In mice, *stella* localizes to all cells during the early stages of preimplantation development and is downregulated from E3.5 prior to inner cell mass formation [16]. Expression of *stella* resumes in the germ cells of the allantoic bud, where it is then found exclusively in migrating germ cells and gonadal primordial germ cells [13, 16]. In the adult ovary, mouse *stella* is then found in oocytes throughout folliculogenesis, as well as in ovulated unfertilized eggs [16]. In mice, *nanog* localizes to all stages of preimplantation embryo development, from fertilized eggs to the inner cell mass of blastocyst [5, 20]. Like the current study in humans, mouse *nanog* is not found in unfertilized eggs. In addition, like mouse *stella*, mouse *nanog* is also expressed in primordial germ cells of the fetal gonad or genital ridge [5], and in the current study, we detected high levels of both *NANOG* and *STELLAR* in human fetal gonads. The difference between previously reported expression of mouse *nanog* and our findings is that we detected *NANOG* in adult testis and ovary in humans, whereas, in mice, there is no apparent expression in any adult tissue except embryonal carcinoma (EC) cells derived from testicular germ cell tumors [5].

Furthermore, in the current study, we detected human *STELLAR* in the adult testis and ovary, whereas, in mice, *stella* is found in mouse ovary only. These differences could be attributed to more sensitive PCR assays used here or to specific differences in expression. It is well documented that *GDF3* is differentially expressed between mice and humans [17, 21]. In particular, in mice, *Gdf3* is found in multiple adult somatic tissues, whereas, in humans, *GDF3* has only been identified in EC cells and is downregulated with differentiation [17]. The relative lack of amino acid identity at the 12p locus between mouse and humans is also interesting, particularly in light of the well-known phenomena that genes

associated with sexual reproduction are more divergent than nonreproductive genes [22].

Although we currently do not know the functional roles of *STELLAR*, *GDF3*, and *NANOG* in gamete formation or differentiation, it is interesting that these three genes, which are germ cell enriched, share the property of amino acid divergence. *OCT-4* has a higher amino acid identity (76%) to its mouse ortholog; however, this is still lower than the majority of human genes (50% of human genes have less than 10% amino acid divergence) [23]. As a consequence, it will be interesting to compare and contrast the functions of these divergent proteins in both mice and humans. The use of mouse and human ES cells may provide a useful model with which to perform this comparative functional analysis. The molecular machinery necessary for establishing the germ cell lineage in humans is almost completely unknown, and the signals necessary for inhibiting a somatic cell fate or promoting a human pluripotent germ cell have yet to be discovered. Our results suggest that genes whose expression is enriched in germ cells or human germ cell tumors and that largely lack expression in somatic tissues could provide a source of novel genes to explore central issues of stem cell pluripotency and self-renewal and germ cell differentiation.

Acknowledgments

We thank the UCSF Fertility Clinic, in particular, *Dr. Shehua Shen*, for obtaining the human oocyte samples. We also thank *Professor Robert N. Taylor, M.D., Ph.D.,* and *Eugene Y. Xu* for helpful comments on the manuscript. This work was supported by grants from the Lance Armstrong Foundation (*P.J.T.*), the Sandler Foundation, and the National Institute of Child Health and Human Development (Grant R01-HD37095) (*R.A.R.P.*).

References

1 Thompson J, Itskovitz-Eldor J, Shapiro S et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145-1147.

2 Hanna LA, Foreman RK, Tarasenko IA et al. Requirement for *Foxd3* in maintaining pluripotent cells of the early mouse embryo. Genes Dev 2002;16:2650-2661.

3 Feldman B, Poueymirou W, Papaioannou VE et al. Requirement of FGF-4 for postimplantation mouse development. Science 1995;267:246-249.

4 Mitsui K, Tokuzawa Y, Itoh H et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 2003;113:631-642.

5 Chambers I, Colby D, Robertson M et al. Functional expression cloning of *Nanog*, a pluripotency sustaining factor in embryonic stem cells. Cell 2003;113:643-655.

6 Matsuda T, Nakamura T, Nakao K et al. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J 1999;15:4261-4269.

7 Rose-John S. GP130 stimulation and the maintenance of stem cells. Trends Biotechnol 2002;20:417-419.

8 Burdon T, Smith A, Savatier P. Signalling cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol 2002;12:432-438.

9 Lin H, Lei JQ, Wininger D et al. Multilineage potential of homozygous stem cells derived from metaphase II oocytes. Stem Cells 2003;21:152-161.

10 Pesce M, Wang X, Wolgemuth DJ et al. Differential expression of the *Oct-4* transcription factor during mouse germ cell differentiation. Mech Dev 1998;71:89-98.

11 de Jong B, Oosterhuis JW, Castedo SM et al. Pathogenesis of adult testicular germ cell tumors. A cytogenic model. Cancer Genet Cytogenet 1990;48:143-167.

12 Murty VV, Dmitrovsky E, Bosl GJ et al. Nonrandom chromosomal abnormalities in testicular and ovarian germ cell tumor cell lines. Cancer Genet Cytogenet 1990;50:67-73.

13 Saitou M, Barton SC, Surani MA. A molecular programme for the specification of germ cell fate in mice. Nature 2002;418:293-300.

14 Pfaffl M, Horgan G, Dempfle L. Relative expression software tool (REST) for groupwise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 2002;30:e36.

15 Moore F, Jaruzelska J, Fox M et al. Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins. Proc Natl Acad Sci USA 2003;100:538-543.

16 Sato M, Kimura T, Kurokawa K et al. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. Mech Dev 2002;113:91-94.

17 Caricasole AA, van Schaik RH, Zeinstra LM et al. Human growth-differentiation factor 3 (hgdf3): developmental regulation in human teratocarcinoma cell lines and expression in primary testicular germ cell tumors. Oncogene 1998;16:95-103.

18 Lee JH, Lee DR, Yoon SJ et al. Expression of *DAZ* (deleted in azoospermia), *DAZL*1 (Daz-like) and protamine-2 in testis and its application for diagnosis of spermatogenic non-obstructive azoospermia. Mol Hum Reprod 1998;4:827-834.

19 Cooke HJ, Lee M, Kerr S et al. A murine homologue of the human *DAZ* gene is autosomal and expressed only in male and female gonads. Hum Mol Genet 1996;5:513-516.

20 Wang S-H, Tsai M-S, Chiang M-F et al. A novel NK-type homeobox gene *ENK* (early embryo specific NK) preferentially expressed in embryonic stem cells. Gene Expr Patterns 2003;3:99-103.

21 McPherron AC, Lee SJ. GDF-3 and GDF-9: two new members of the transforming growth factor-beta superfamily containing a novel pattern of cysteines. J Biol Chem 1993;268:3444-3449.

22 Swanson W, Vacquier V. The rapid evolution of reproductive proteins. Nat Rev Genet 2002;3:137-144.

23 Makalowski W, Boguski M. Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. Proc Natl Acad Sci USA 1998;95:9407-9412.