STEM CELLS®
Original Article

Properties of Pluripotent Human Embryonic Stem Cells BG01 and BG02

XIANMIN ZENG,[a] TAKUMI MIURA,[b] YONGQUAN LUO,[b] BHASKAR BHATTACHARYA,[c] BRIAN CONDIE,[d] JIA CHEN,[a] IRENE GINIS,[b] IAN LYONS,[d] JOSEF MEJIDO,[c] RAJ K. PURI,[c] MAHENDRA S. RAO,[b] WILLIAM J. FREED[a]

[a]Cellular Neurobiology Research Branch, National Institute on Drug Abuse, Department of Health and Human Services (DHHS), Baltimore, Maryland, USA; [b]Laboratory of Neuroscience, National Institute of Aging, DHHS, Baltimore, Maryland, USA; [c]Laboratory of Molecular Tumor Biology, Division of Cellular and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland, USA; [d]BresaGen Inc., Athens, Georgia, USA

Key Words. *Embryonic stem cells · Differentiation · Microarray*

ABSTRACT

Human ES (hES) cell lines have only recently been generated, and differences between human and mouse ES cells have been identified. In this manuscript we describe the properties of two human ES cell lines, BG01 and BG02. By immunocytochemistry and reverse transcription polymerase chain reaction, undifferentiated cells expressed markers that are characteristic of ES cells, including SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, and OCT-3/4. Both cell lines were readily maintained in an undifferentiated state and could differentiate into cells of all three germ layers, as determined by expression of β-tubulin III neuron-specific molecule (ectoderm), cardiac troponin I (cardiomyocytes, mesoderm), and α-fetoprotein (endoderm). A large-scale microarray (16,659 genes) analysis identified 373 genes that were expressed at three-fold or higher levels in undifferentiated BG01 and BG02 cells as compared with pooled human RNA. Ninety-two of these genes were also highly expressed in four other hES lines (TE05, GE01, GE09, and pooled samples derived from GE01, GE09, and GE07). Included in the list are genes involved in cell signaling and development, metabolism, transcription regulation, and many hypothetical proteins. Two focused arrays designed to examine transcripts associated with stem cells and with the transforming growth factor-β superfamily were employed to examine differentially expressed genes. Several growth factors, receptors, and components of signaling pathways that regulate embryonic development, in particular the nodal signaling pathway, were detected in both BG01 and BG02. These data provide a detailed characterization and an initial gene expression profile for the BG01 and BG02 human ES cell lines. ***Stem Cells*** *2004;22:292-312*

INTRODUCTION

Embryonic stem (ES) cells, isolated from the blastocysts of preimplantation embryos, are pluripotent and have the capability to generate all the differentiating cells present in the embryo. ES cells were first described in mice and recently have been identified from multiple species including

Correspondence: Xianmin Zeng, Ph.D., Development and Plasticity Section, Cellular Neurobiology Research Branch, National Institute on Drug Abuse, 333 Cassell Drive, Baltimore, Maryland 21224, USA. Telephone: 410-550-6565 (ext 138); Fax: 410-550-1621; e-mail: xzeng@intra.nida.nih.gov Mahendra S. Rao, Ph.D., Laboratory of Neuroscience, National Institute of Aging, 33 Cassell Drive, Baltimore, Maryland 21224, USA. Telephone: 410-558-8204; Fax: 410-558-8249; e-mail: raomah@grc.nia.nih.gov
Received August 22, 2003; accepted for publication October 28, 2003.

subhuman primates (rhesus and marmoset) and human [1-4]. Because of their unique properties, human (h)ES cells could be used for repair and replacement of cells or tissues lost due to disease or trauma. A universal bank of well-characterized hES cells from which specific cells can be generated would potentially be invaluable for basic research and cell replacement therapy.

To date, 78 different lines from the National Institutes of Health (NIH) registry have been identified and tentatively classified as ES cells based on general morphological similarity. Early experiments suggest that the properties of hES cells differ in some respects from mouse ES cells [4]. Eleven of these lines are currently available for research purposes, and limited data on the biology of 26 of these lines are available [5]. Two of these lines, BG01 and BG02, were generated from embryos whose poor development was such that in the course of usual IVF practice they would have been discarded because 6-7 days post fertilization, fully 1-2 days after the usual time of embryo transfer, they had not developed sufficiently to survive cryopreservation. A report of their basic biology is available [6]; however, there are no published data on molecular characterization of these cell lines.

Considerable additional information is, however, available on other hES cell lines, notably those derived by *Thomson* and colleagues [4]. Teratocarcinoma formation, long-term stability, derivation of feeder-free subclones, microarray analysis, genome scan, and serial analysis of gene expression analysis have been completed or initiated for several lines [7-11]. Genetic modification, including homologous recombination, has been reported [12, 13]. Differentiation into multiple phenotypes, including cardiac, hepatic, pancreatic, neural, and hematopoietic lineages, has been described [14-21]. Whether BG01 and BG02 lines have similar properties remains to be determined.

It is unlikely that all hES cell lines will be identical or equally stable in culture. Some differences in human cell lines have been described [5], although whether they reflect differences in the methods of propagation or illustrate underlying differences in biology remains to be determined. We also note that when rodent ES lines have been examined, strain differences in isolation and propagation of lines have been described [22, 23]. Indeed, uniformly successful isolation is not possible in all mouse strains, and it has been difficult to generate rat ES cell lines [24].

In an effort to understand the properties of the BG lines, we have propagated BG01 and BG02 in culture and examined their growth; differentiation characteristics; and gene expression patterns using immunostaining, reverse transcriptase polymerase chain reaction (RT-PCR), and microarray analysis. We show that BG01 and BG02 have the capacity to differentiate into cells that express divergent tissue-specific antigens consistent with pluripotency and express markers similar to other pluripotent hES cells. No significant differences were observed in gene expression profiles between these two lines.

Materials and Methods

Isolation and Growth of ES Cells

hES cell lines BG01 and BG02 were obtained from BresaGen (Athens, GA) and cultured according to manufacturer instructions. ES cells were maintained on mitomycin-C-inactivated mouse embryonic fibroblast (MEF, from strain SVB, 1×10^6 cells/35 mm dish) feeder cells in Dulbecco's-modified Eagle's medium/Ham's F12 (1:1) supplemented with 15% fetal bovine serum (FBS), 5% knockout serum replacement (KSR), 2 mM nonessential amino acids, 2 mM L-glutamine, 50 μg/ml Penn-Strep (all from Invitrogen; Carlsbad, CA; http://www.invitrogen.com), 0.1 mM β-mercaptoethanol (Specialty Media; Phillipsburg, NJ; http://www.specialtymedia.com), and 4 ng/ml of basic fibroblast growth factor (bFGF; Sigma; St. Louis, MO; http://www.sigma aldrich. com). Cells were passaged by incubation in cell dissociation buffer or trypsin (Invitrogen), dissociated, and then seeded at about 20,000 cells/cm^2. Under such culture conditions, the ES cells were passaged every 4-5 days. For freezing, cells were resuspended in medium containing 25% FBS, 65% hES medium, and 10% dimethylsulfoxide at 1×10^6 cells/ml at approximately 1°C per minute.

Differentiation In Vitro

ES cell cultures were dissociated into small clumps by collagenase IV (Sigma) by incubating at 37°C for 5 minutes. The hES cell colonies were pelleted, resuspended in hES medium without bFGF (differentiation medium), and cultured in 6-well plates for 7 days with a medium change every second day. ES cell colonies grew in suspension as embryoid bodies (EBs), while remaining feeder cells adhered to the plate. The EBs were transferred into a new plate and were further cultured for 7 days before immunostaining.

Immunocytochemistry

Expression of stem cell markers was examined by immunocytochemistry, and staining procedures were as described previously [25]. Briefly, the ES cells were fixed with 4% paraformaldehyde and permeabilized with 0.1% Triton X-100. After blocking, the cells were incubated with primary antibody. The primary antibodies and the dilutions used are stage-specific embryonic antigen (SSEA)-1, SSEA-3, and SSEA-4 (Developmental Studies Hybridoma Bank,

University of Iowa; http://www.uiowa.edu; 1:50); tumor recognition antigen (TRA)-1-60 and TRA-1-81 (Chemicon; Temecula, CA; http://www.chemicon.com; 1:100); octamer-binding transcription factor (OCT)-3/4 and cTnI (Santa Cruz Biotechnology; Santa Cruz, CA; http://www.scbt.com, 1:100); TuJ1 and α-fetoprotein (Sigma; 1:2000); smooth muscle actin (Sigma; 1:200); and nestin (BD Biosciences; San Jose, CA, http://www.bdbiosciences.com; 1:500). Localization of antigens was visualized by using appropriate secondary antibodies (Alexa fluor 594 or 488, Molecular Probes; Eugene, OR; http://www.probes.com).

RT-PCR Analysis

Total RNA was isolated with TRIzol (Invitrogen), a modification of the guanidine isothiocyanate-phenol-chloroform extraction method. The undifferentiated state of cultured cells was verified by immunostaining of ES markers such as OCT-3/4, SSEA-4, and TRA-1-60. cDNA was synthesized using 2.5 μg total RNA in a 20-μl reaction with Superscript II (Invitrogen) and oligo $(dT)_{12\text{-}18}$ (Promega; Madison, WI; http://www.promega.com). One microliter RNase H (Invitrogen) was added to each tube and incubated for 20 minutes at 37°C before proceeding to the RT-PCR analysis. The PCR primers for OCT-3/4, SOX-2, REX-1, UTF1, hTERT, Dppa5, Cx43, Cx45, and ABCG2 were described by *Ginis et al*. [26]; the primers for Nanog, FLJ13072, KIAA1265, MGC27165, ZNF342, DNMT3L, DAX-1, Eras, TUBB5, KRT8, KRT18, cardiac actin, and galanin are listed in Table 1.

For each PCR reaction, 0.5 ml of 1:10 diluted cDNA template was used in a 50-ml reaction volume with the Taq DNA polymerase (Invitrogen). The cycling parameters were 94°C, 1 minute; 58°C, 1 minute; and 72°C, 1 minute, for 30 cycles. The PCR cycle was preceded by an initial denaturation of 3 minutes at 94°C and followed by a final extension of 10 minutes at 72°C.

Large-Scale Oligonucleotide Microarray

The microarray used in this analysis contained 16,659 70-bp oligonucleotides designed from 750 bases of the 3′ end of each open reading frame that represents the largest verified set available. Twenty micrograms of total RNA from each of BG01 and BG02 and universal human RNA (huRNA, Clontech; Palo Alto, CA; http://www.clontech.com) were labeled with Cy5 and Cy3, respectively, and duplicate arrays were hybridized and processed using the modified procedure developed at the Center for Biologics Evaluation and Research (CBER) microarray program under an interagency agreement between CBER and the Advanced Technology Center of the National Cancer Institute. The images were captured under wavelengths appropriate for both Cy3 and Cy5 photomultiplier tubes (PMTs; 600V-750V) using a GenePix 4000B scanner (Axon Instruments, Inc.; Union City, CA; http://www.axon.com). At this PMT range, the images gave the best signal-to-noise ratio. The data were initially analyzed using Gene Pix software. Both the scanned image and analyzed data files were uploaded into the Center for Information microarray database (mAdb) (http://www.nciarray.nci.nih.gov). The data were analyzed using two different analytic tools: single array viewer tools and extended data extraction tool. Scatter plot analysis of a control array with Cy5- and Cy3-labeled total RNA derived from huRNA showed equal hybridization of spots as indicated by a straight line from X to Y axis, indicating good quality of most spots and uniform hybridization (not shown).

Focused Microarray Analysis

The nonradioactive GEArray Q series cDNA expression array filters for human stem cell genes and transforming growth factor (TGF) β/BMPl pathway genes (Hs601 and Hs023; SuperArray Inc.; Frederick, MD; http://www.superarray.com) [26] were used according to the manufacturer's protocol. The biotin dUTP-labeled cDNA probes were specifically generated in the presence of a designed set of gene-specific primers using total RNA (4 mg/filter) and 200 U Moloney murine leukemia virus reverse transcriptase (Promega). The array filters were hybridized with biotin-labeled probes at 60°C for 17 hours. After that, the filters were washed twice with 2 × SSC/1% SDS and then twice with 0.1 × SSC/1% SDS at 60°C for 15 minutes each. Chemiluminescent detection steps were performed by incubation of the filters with alkaline phosphatase-conjugated streptavidin and CDP-Star substrate. Array membranes were exposed to x-ray film. Quantification of gene expression on the array was performed with ScionImage software. cDNA microarray experiments were done twice with new filters and RNA isolated at different times. Results from the focused array were independently confirmed, and the array itself was validated using procedures previously described [27].

Results

Growth and Morphology

BG01 and BG02 were cultured on MEF feeders and grew as colonies of tightly compacted undifferentiated cells resembling mouse and hES cells (Fig. 1). Like murine ES cells, BG01 and BG02 colonies had a high nuclear/cytoplasmic ratio. These cells were propagated in vitro for more than 40 passages and maintained a normal karyotype [6]. Successful passage of cells was achieved by using either trypsin or nonenzymatic cell dispersal buffers, and we found that both worked equally well. In general, the cultures

Table 1. PCR primer pairs used in this study

Gene	Size	Primer	Sequence (5′ to 3′)
ABCG2	684 bp	hABCG2-F	GTTTATCCGTGGTGTGTCTGG
		hABCG2-R	CTGAGCTATAGAGGCCTGGG
Cardiac actin	164 bp	hCRD-F	GCCCTGGATTTTGAGAATGA
		hCRD-R	ATGCCAGCAGATTCCATACC
Connexin 43	295 bp	hCx43-F	TACCATGCGACCAGTGGTGCGCT
		hCx43-R	GAATTCTGGTTATCATCGGGGAA
Connexin 45	819 bp	hCx45-F	CTATGCAATGCGCTGGAAACAACA
		hCx45-R	CCCTGATTTGCTACTGGCAGT
DNA methyltransferase 3-like (DNMT3L)	189 bp	hDNMT3L-F	CTCTCAAGCTCCGTTTCACC
		hDNMT3L-R	GTACAGGAAGAGGGCATCCA
Dppa 5	353 bp	hDppa5-F	ATGGGAACTCTCCCGGCACG
		hDppa5-R	TCACTTCATCCAAGGGCCTA
E-Ras	152 bp	hERas-F	GCTGTCTGTGATGGTGTGCT
		hERas-R	TCTCCAGCAGTGGTCACAAG
Hypothetical protein FLJ13072	179 bp	hFLJ-F	TCTTCCGCAGACCAGTACCT
		hFLJ-R	TCTCGGCAGAAGTCAATGTG
pro-Galanin	156 bp	hGal-F	AAGGAAAAACGAGGCTGGAC
		hGal-R	GGACCTGTCAAAGCTTCCTG
Glyceraldehyde-3-phosphate dehydrogenase	474 bp	hGAPD-F	GCTCAGACACCATGGGGAAGGT
		hGAPD-R	GTGGTGCAGGAGGCATTGCTGA
Hypothetical protein KIAA1265	194 bp	hKIAA-F	CCTTTGCCCTGCATTGTTAT
		hKIAA-R	AGATCACGCCTAGCAAGGAA
Keratin 8	161 bp	hKRT8-F	TGAGGTCAAGGCACAGTACG
		hKRT8-R	TGATGTTCCGGTTCATCTCA
Keratin 18	164 bp	hKRT18-F	CACAGTCTGCTGAGGTTGGA
		hKRT18-R	GAGCTGCTCCATCTGTAGGG
Hypothetical protein MGC27165	216 bp	hMGC-F	TTGGTCCCTGGCTAATTCAC
		hMGC-R	TTGTGCATCTTCTGGCTTTG
Nanog (Hypothetical protein FLJ12581)	158 bp	hNanog-F	CAAAGGCAAACAACCCACTT
		hNanog-R	TCTGCTGGAGGCTGAGGTAT
NR0B1 (Dax-1 protein)	154 bp	hDax1-F	GACTCCAGTGGGGAACTCAG
		hDax1-R	ATGATGGGCCTGAAGAACAG
Oct-3/4	171 bp	hOct3/4-F	CTTGCTGCAGAAGTGGGTGGAGGAA
		hOct3/4-R	CTGCAGTGTGGGTTTCGGGCA
Rex-1	559 bp	hRex1-F	TGAAAGCCCACATCCTAACG
		hRex1-R	CAAGCTATCCTCCTGCTTTGG
Sox-2	437 bp	hSox2-F	ATGCACCGCTACGACGTGA
		hSox2-R	CTTTTGCACCCCTCCCATTT
hTERT	147 bp	LT5	CGGAAGAGTGTCTGGAGCAA
		LT6	GGATGAAGCGGAGTCTGGA
β-tubulin 5	161 bp	hTUBB5-F	CAGTGACCTGCAACTGGAGA
		hTUBB5-R	GATTGGCCAAACACGAAGTT
UTF-1	230 bp	hUTF1-F	ACCAGCTGCTGACCTTGAAC
		hUTF1-R	TTGAACGTACCCAAGAACGA
Zinc finger protein 342	193 bp	hZNF342-F	GAAGGCATCACCCAAAAAGA
		hZNF342-R	GCGGTTGAGCTTACTGCTCT

became confluent 4-5 days after seeding and the cells had a doubling time of 30-35 hours, similar to that reported for other hES cells grown on feeders or feeder-free conditioned medium. In addition, these cells recovered quickly from

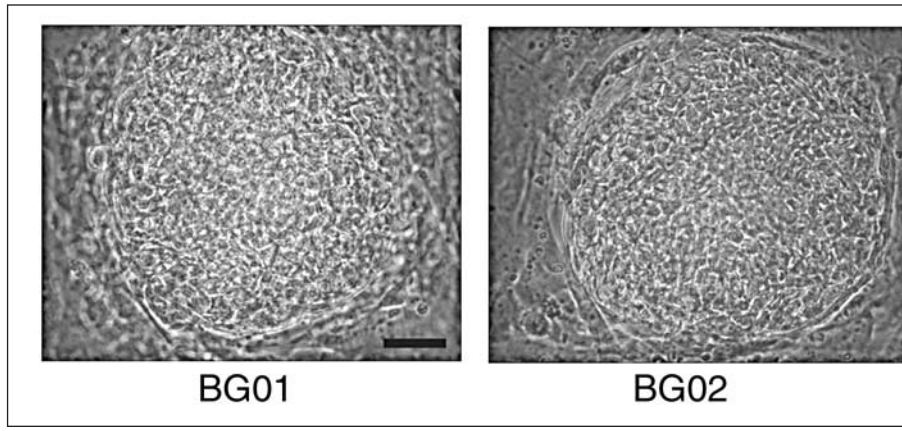

Figure 1. Morphology of undifferentiated BG01 and BG02 cells. *Phase contrast microscopy of undifferentiated BG01 and BG02 cells grown on an MEF feeder layer. The scale bar is 20 μm.*

frozen vials within 2-3 days, and an undifferentiated state was easily and reliably maintained with no instances of spontaneous differentiation. Spontaneous differentiation was, however, observed when BG01 or BG02 cells were seeded on MEFs at lower density (e.g., 1×10^5 cells/35 mm dish).

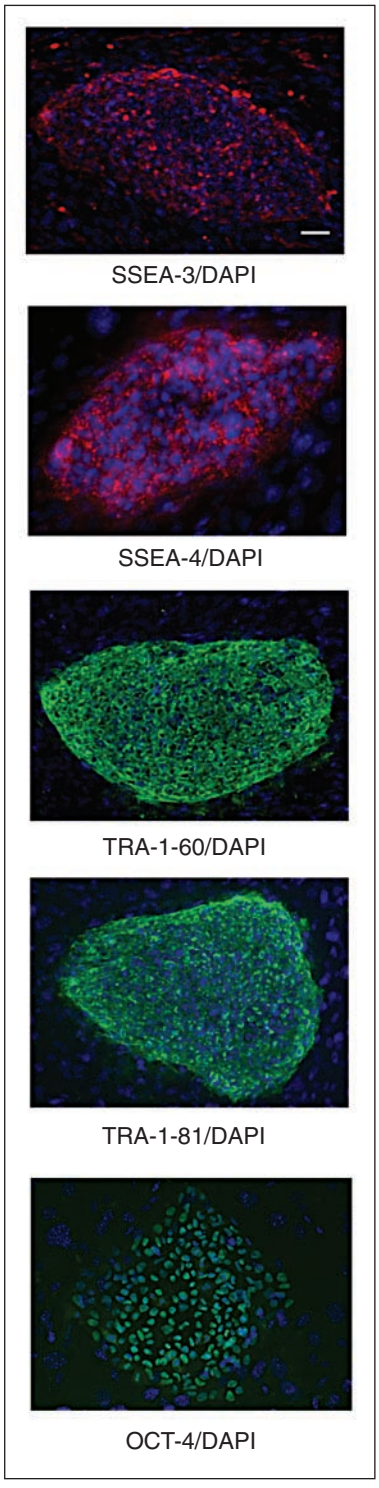

Figure 2. Morphology of undifferentiated BG01 and BG02 cells and expression of markers by immunocytochemistry. Both BG01 and BG02 are strongly positive for SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, and OCT-3/4. Antibody staining is in red (SSEA-3 and SSEA-4) or green (TRA-1-60, TRA-1-81, and OCT-3/4), while nuclear DAPI staining is in blue. The scale bar is 20 µm for SSEA-4 and OCT-3/4 and 10 µm for the others.

Cell Surface and ES Cell Marker Expression

Immunocytochemistry was used to analyze whether BG01 and BG02 were similar to other hES cell lines in expressing cell surface markers that characterize undifferentiated pluripotent stem cells. These include SSEA-3 and SSEA-4; tumor recognition antigen, TRA-1-60 and TRA-1-81; and the POU transcription factor, OCT-3/4.

Undifferentiated BG01 and BG02 cells were strongly positive for TRA-1-60, TRA-1-81, SSEA-4, and OCT-3/4, but negative for SSEA-1. Almost all of the colonies were positive for TRA-1-60, TRA-1-81, SSEA-3, SSEA-4, and OCT-3/4, and the majority of cells in the colonies were stained for these markers (Fig. 2). Almost no positive staining was observed outside the ES colonies or in the feeder cells for TRA-1-60, TRA-1-81, SSEA-4, and OCT-3/4, but some positive cells were found outside the colonies for SSEA-3 (Fig. 2). Staining intensity for SSEA-4, TRA-1-60, TRA-1-81, and OCT-3/4 was consistently strong both within individual colonies and among the colonies, but staining intensity was weaker for SSEA-3.

RT-PCR was used to confirm some of the markers analyzed by immunochemistry and to examine additional markers that are associated with stem cells. Primers that span intron-exon boundaries for the TERT gene were used to assess possible genomic contamination. As expected, messages for telomerase reverse transcriptase TERT, homeobox-domain transcription factor SOX-2, zinc finger protein REX-1, and a gene similar to developmental pluripotency-associated 5 (Dppa 5) were detected in undifferentiated BG01 and BG02 cells (Fig. 3). Transcriptional coactivator UTF1 was also expressed in BG01 and BG02. Transcripts for some cell surface markers reported for blastocysts or stem cells, such as the gap junction proteins connexin-43 and connexins-45 (Cx-43 and Cx-45), and ABC transporter ABCG were also detected in undifferentiated BG01 and BG02 cells (Fig. 3). Table 2 summarizes the markers expressed by BG01 and BG02 using either immunocytochemistry, RT-PCR, or both.

Ability to Differentiate into Cell Types from All Three Germ Layers

The capacity of BG01 and BG02 to differentiate in vitro was evaluated. Differentiation of hES cells was initiated by culturing in the absence of an MEF feeder layer and bFGF. Under these conditions, hES cells formed embryoid

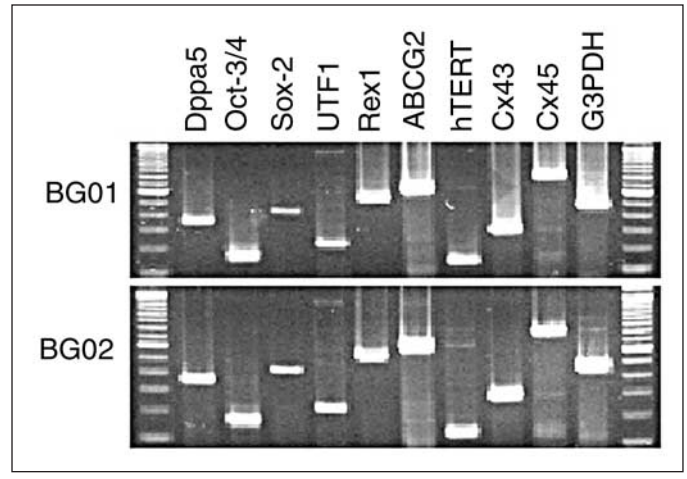

Figure 3. Expression of markers associated with ES cells in BG01 and BG02 by RT-PCR. Genes known to be associated with the pluripotent state (hTERT, OCT-3/4, SOX-2, and REX-1) are expressed in both BG01 and BG02; five additional genes, Dppa5, UTF1, ABCG2, Cx43, *and* Cx45 *are also expressed by BG01 and BG02. PCR was performed using gene-specific primers (Table 1) with glyceraldehyde-3-phosphate dehydrogenase as a control. Markers are a 100-bp ladder with the lowest band being 100 bp.*

Table 2. Expression of markers is associated with stem cells in undifferentiated BG01 and BG02 cells

Marker	BG01 RT-PCR	BG01 Staining	BG02 RT-PCR	BG02 Staining
SSEA-1		–		–
SSEA-3		+		+
SSEA-4		+		+
TRA-1-60		+		+
TRA-1-81		+		+
TERT	+		+	
OCT-4	+	+	+	+
SOX-2	+		+	
REX-1	+		+	
UTF1	+		+	
Dppa5	+		+	
Cx43	+		+	
Cx45	+		+	
ABCG2	+		+	

bodies (EBs) of heterogeneous cells. TuJ1-positive neurons, cardiac troponin I-positive cardiomyocytes, and α-fetoprotein-positive cells (Fig. 4) were identified in EBs by immunocytochemistry. In addition, cells positive for nestin and smooth muscle actin were identified (Fig. 4). This result indicated that BG01 and BG02 were able to differentiate in vitro into ectodermal, mesodermal, and endodermal derivatives.

Microarray Analysis of Gene Expression in Undifferentiated Cells

The large-scale oligonucleotide microarray was employed to examine the overall gene expression profile of undifferentiated BG01 and BG02 cells. The probes used for this array included 1,987 hypothetical proteins and 72 expressed sequence tags (ESTs) and span approximately 50% of the human genome. RNA from undifferentiated BG01 and BG02 cells was compared with pooled huRNA that hybridizes with most genes on the array. The array results are shown in Figure 5. Figure 5A and B show the image profiles of BG01 and BG02, Figure 5C and D show scatter plot analyses of BG01 and BG02, and Figure 5E shows a comparison of highly expressed genes (threefold or higher level) among BG01, BG02, and other ES cell lines.

A total of 373 genes common to both BG01 and BG02 were identified as being differentially expressed, defined as a threefold or greater increase as compared with huRNA (Table 3). Among them, 92 genes, including several stemness genes known to be expressed in mES or hES cells such as OCT-3/4, NANOG, Cripto/TDGF1, Cx43, and galanin, are also differentially expressed in an additional four hES lines (TE06, GE01, GE09, and a pooled set of subclones derived from GE01, GE09, and GE07). Expression of several of these 92 genes in BG01 and BG02 was confirmed by RT-PCR (Fig. 6). Nanog (a recently identified hypothetical protein FLJ12581 critical for maintaining pluripotency of mES cells) and three other hypothetical proteins, FLJ13072, KIAA1265, and MGC27165, were all expressed in both BG01 and BG02 (Fig. 6A). ZNF342, DNMT3L, DAX-1, and Eras transcripts were detected by RT-PCR (Fig. 6B).

The remaining 281 of the 373 genes that were found to be differentially expressed by threefold or greater in BG01 and BG02 were not found to be differentially expressed in the other hES cell lines. These included metabolic-related genes, ribosomal proteins, histone proteins, and many hypothetical proteins. In addition, 19 novel genes were also

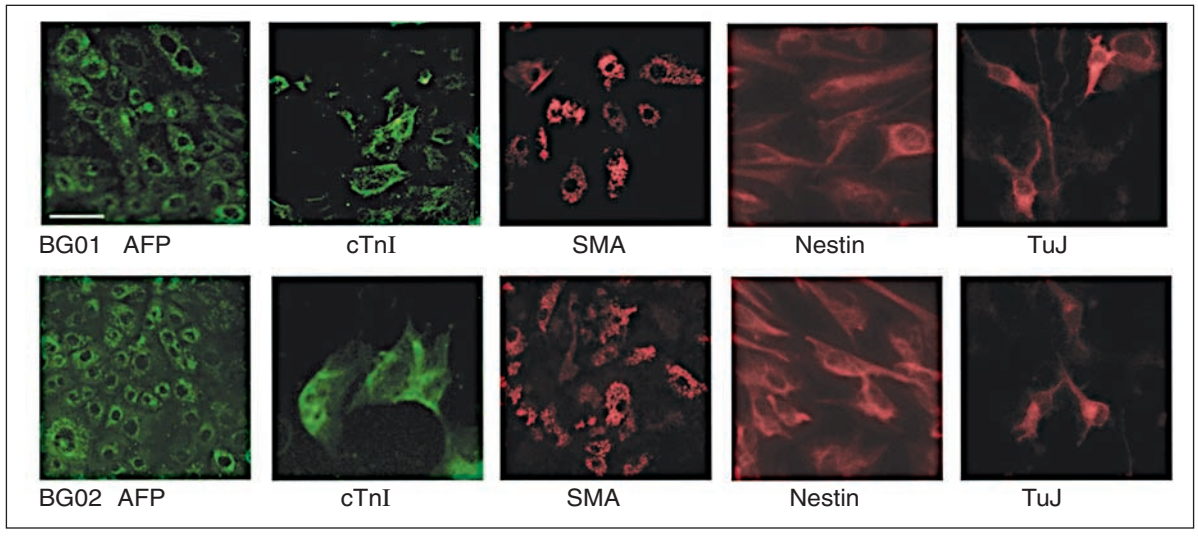

Figure 4. In vitro differentiation of BG01 and BG02 via embryoid bodies. *Differentiation of hES cells was initiated by forming EBs in the absence of MEFs and bFGF. Positive immunostaining was identified for AFP, cTnI, SMA, nestin, and TuJ1, indicating that both lines can differentiate to express markers of ectoderm, mesoderm, and endoderm. The scale bar is 20 μm.*

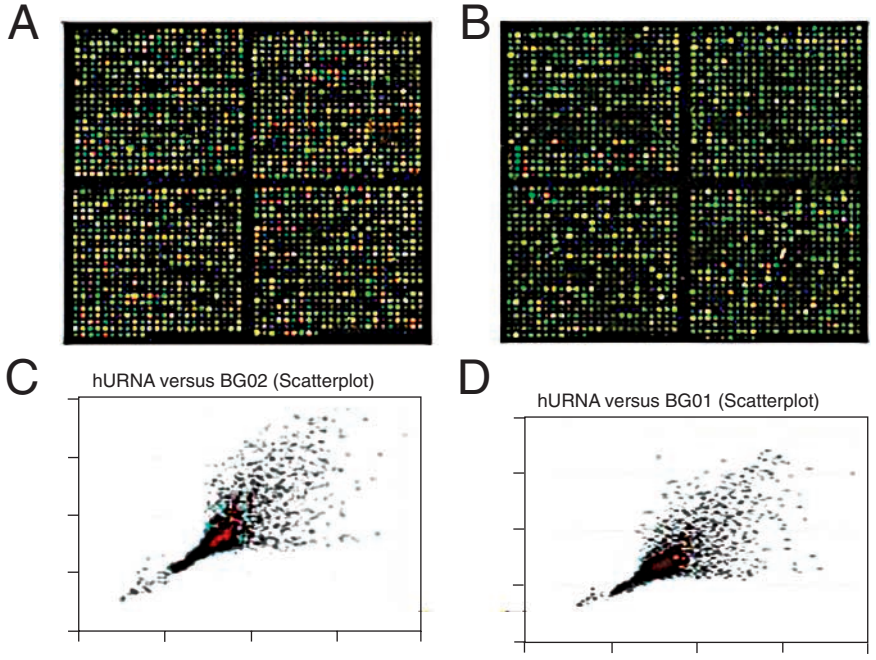

Figure 5. Expression profiling of BG01 and BG02 hES cell lines. *A and B): Representative images after competitive hybridization of oligonucleotide arrays using cy5-labeled RNA from BG01 (A) and BG02 (B) hES cell lines and cy3-labeled huRNA. Total RNA (5 µg) isolated from hES cell lines and huRNA was labeled with cy5 and cy3 dyes, respectively, and used to hybridize oligonucleotide arrays containing ~17,000 features. The images were collected using a GenePix scanner. C and D): Scatter plot analysis of cy5-labeled genes in BG01 (C) and BG02 (D) and cy3-labeled genes in hURNA to show differential gene expression. E). Venn diagram showing similarity of highly expressed genes (threefold and higher) in BG01 and BG02 cell lines at 99% confidence interval. A total of 1,006 genes were highly expressed in BG01, 406 were highly expressed in BG02, and 373 of these were common to both cell lines. Of the 373 transcripts common to BG01 and BG02, 92 were also found to be highly expressed in other hES cell lines and 281 were found only in BG01 and BG02.*

identified and are listed separately (Table 4). We noted that a few genes that are considered as markers of differentiation were expressed at high levels in both undifferentiated BG01 and BG02 cells. These included keratin 8, 14, and 19; cardiac actin; and tubulin. Expression of these genes in undifferentiated ES cells was also found by RT-PCR (Fig. 6C).

Nodal Signaling

Among the genes that were differentially expressed are the TGF-β superfamily member Nodal and its antagonist Lefty. Nodal signaling is known to play an important role in endoderm formation, early embryonic patterning, and left-right axis positioning. To further examine expression of genes related to the Nodal signaling pathway in undifferentiated cells, we employed a focused microarray (96 genes) containing probes of the TGF-β superfamily and key proteins involved in the TGF-β signal-transduction pathway. Figure 7A shows a hybridization profile of BG01 and BG02, and Figure 7B summarizes the genes detected in both BG01 and BG02.

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold)

Gene	Description	UniGene	BG02	BG01
GAL	galanin	Hs.278959	32.7297	30.6676
LIN-28	RNA-binding protein LIN-28	Hs.86154	27.7381	20.2924
POU5F1	POU domain, class 5, transcription factor 1	Hs.2860	27.6215	15.1775
SEMA6A	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A	Hs.263395	25.1226	13.9032
FLJ12505	hypothetical protein FLJ12505	Hs.125741	23.5175	16.8017
NPM1	nucleophosmin (nucleolar phosphoprotein B23, numatrin)	Hs.355719	23.3351	14.0028
KRT18	keratin 18	Hs.406013	22.9837	12.7228
SNRPF	small nuclear ribonucleoprotein polypeptide F	Hs.105465	22.7924	5.2743
ACTC	actin, alpha, cardiac muscle	Hs.118127	22.2086	26.4846
SSB	Sjogren syndrome antigen B (autoantigen La)	Hs.83715	19.6682	9.5376
DLG7	discs, large homolog 7 (Drosophila)	Hs.77695	19.2045	7.7198
SLC16A1	solute carrier family 16 (monocarboxylic acid transporters), member 1	Hs.75231	19.2034	10.0484

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	Hs.211556	19.0097	9.6355
CRABP2	cellular retinoic acid binding protein 2	Hs.183650	18.5854	18.7416
	Human HL14 gene encoding beta-galactoside-binding lectin, 3′ end, clone 2		18.1754	19.5927
EIF4A1	eukaryotic translation initiation factor 4A, isoform 1	Hs.356129	18.1336	8.3636
LEFTB	left-right determination, factor B	Hs.278239	18.0146	8.4236
CST4	cystatin S	Hs.56319	17.9819	4.0157
FABP5	fatty acid binding protein 5 (psoriasis-associated)	Hs.408061	17.8858	9.4236
CYP26A1	cytochrome P450, family 26, subfamily A, polypeptide 1	Hs.150595	17.8454	5.4796
	Human DNA sequence from clone RP11-248N6 on chromosome 13.		17.374	12.933
LECT1	leukocyte cell derived chemotaxin 1	Hs.97932	17.3608	7.3421
Nup37	nucleoporin Nup37	Hs.5152	17.1315	6.0084
DKFZP586L072	DKFZP586L0724 protein	Hs.26761	17.0567	7.5128
RPL6	ribosomal protein L6	Hs.409045	16.8328	6.7963
KRT8	keratin 8	Hs.242463	16.5164	10.2776
GPC4	glypican 4	Hs.58367	15.5943	7.2993
LDHB	lactate dehydrogenase B	Hs.234489	15.2657	5.63
RPS24	ribosomal protein S24	Hs.180450	14.756	5.6245
DKC1	dyskeratosis congenita 1, dyskerin	Hs.4747	14.5261	4.5643
PODXL	podocalyxin-like	Hs.16426	14.3853	11.2914
SMS	spermine synthase	Hs.89718	14.3544	6.0537
	Human DNA sequence from clone RP11-392A19 on chromosome 13.		14.1975	6.539
SPS	selenophosphate synthetase	Hs.124027	13.9526	9.9908
NME1	non-metastatic cells 1, protein (NM23A)	Hs.118638	13.6861	6.6279
DDX21	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 21	Hs.169531	13.536	5.9288
PDHB	pyruvate dehydrogenase (lipoamide) beta	Hs.979	13.5264	4.9302
HSPCA	heat shock 90kDa protein 1, alpha	Hs.356531	13.4254	7.5954
HUMAUANTIG	nucleolar GTPase	Hs.75528	13.3235	4.9662
GDF3	growth differentiation factor 3	Hs.86232	12.6755	8.1771
	Human DNA sequence from clone RP11-438F9 on chromosome 13.		12.617	6.9879
NUP107	nuclear pore complex protein	Hs.236204	12.588	5.1862
HNRPAB	heterogeneous nuclear ribonucleoprotein A/B	Hs.81361	12.3917	7.5878
SSBP1	single-stranded DNA binding protein	Hs.923	12.3331	5.6571
CCT2	chaperonin containing TCP1, subunit 2 (beta)	Hs.432970	12.2648	4.877
	Human DNA sequence from clone 167F1 on chromosome 6p22.1-22.3. Contains a KRT18 (Keratin, type 1 Cyt)		12.2497	4.5162
HMGIY	high-mobility group (nonhistone chromosomal) protein isoforms I and Y		12.1465	13.1788
CCT8	chaperonin containing TCP1, subunit 8 (theta)	Hs.15071	12.1346	6.7962
RBBP6	retinoblastoma binding protein 6	Hs.91065	11.4926	5.2396
KIF4A	kinesin family member 4A	Hs.279766	11.2156	6.5153
PIR51	RAD51-interacting protein	Hs.24596	11.0008	3.3822
TRA1	tumor rejection antigen (gp96) 1	Hs.82689	10.989	4.4254
C1GALT1	core 1 UDP-galactose:N-acetylgalactosamine-alpha-R beta 1,3-galactosyltransferase	Hs.46744	10.9054	4.3861
KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Hs.159557	10.8785	7.32
LDHA	lactate dehydrogenase A	Hs.2795	10.6996	8.6252
UGP2	UDP-glucose pyrophosphorylase 2	Hs.77837	10.6941	6.2481

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
PGK1	phosphoglycerate kinase 1	Hs.78771	10.6804	5.4813
CCT4	chaperonin containing TCP1, subunit 4 (delta)	Hs.79150	10.6741	5.645
NME2	non-metastatic cells 2, protein (NM23B)	Hs.433416	10.6292	7.9767
DKFZP564O046	DKFZP564O0463 protein	Hs.273344	10.4118	3.9574
ABCE1	ATP-binding cassette, sub-family E (OABP), member 1	Hs.12013	10.3581	6.1063
HSPD1	heat shock 60kDa protein 1 (chaperonin)	Hs.79037	10.3413	5.49
HSPCB	heat shock 90kDa protein 1, beta	Hs.74335	10.2904	5.7852
TBCA	tubulin-specific	Hs.433254	10.2576	3.9957
KPNB1	karyopherin (importin) beta 1	Hs.180446	10.2516	5.2268
SAS10	disrupter of silencing 10	Hs.322901	10.2158	3.0576
SNRPA1	small nuclear ribonucleoprotein polypeptide A′	Hs.80506	10.1138	5.3005
	Human DNA sequence from clone RP3-447F3 on chromosome 20.		10.0873	7.0067
SLC38A1	solute carrier family 38, member 1	Hs.18272	9.9938	5.2499
PSMA2	proteasome (prosome, macropain) subunit, alpha type, 2	Hs.411773	9.9558	6.847
NASP	nuclear autoantigenic sperm protein (histone-binding)	Hs.380400	9.7835	7.2138
NUDT5	nudix (nucleoside diphosphate linked moiety X)-type motif 5	Hs.301957	9.7082	6.7425
CCNB1	cyclin B1	Hs.23960	9.6941	9.3155
PPAT	phosphoribosyl pyrophosphate amidotransferase	Hs.311	9.6117	3.728
GJA1	gap junction protein, alpha 1, 43kDa (connexin 43)	Hs.74471	9.5191	15.2909
PRO2013	hypothetical protein PRO2013	Hs.238205	9.4974	4.537
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen, 80kDa)	Hs.84981	9.4109	4.0437
TD-60	RCC1-like	Hs.284146	9.3161	5.5608
CDC2	cell division cycle 2, G_1 to S and G_2 to M	Hs.334562	9.2409	6.5016
NS	nucleostemin	Hs.279923	9.201	5.3758
PWP1	nuclear phosphoprotein similar to S. cerevisiae PWP1	Hs.172589	9.1489	3.7634
FLJ21908	hypothetical protein FLJ21908	Hs.26750	9.1163	3.8205
RPL23A	ribosomal protein L23a	Hs.406616	9.0718	4.3649
CCNC	cyclin C	Hs.118442	9.0591	4.5857
HCA66	hepatocellular carcinoma-associated antigen 66	Hs.30670	9.0432	3.8016
LMNB1	lamin B1	Hs.89497	8.9928	3.0784
MRPL1	mitochondrial ribosomal protein L1	Hs.283693	8.9669	5.0821
KIAA1573	KIAA1573 protein	Hs.24790	8.9336	5.9659
SFRS1	splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor)	Hs.73737	8.8251	4.3408
	Homo sapiens hypothetical gene supported by XM_010918 (LOC65365), mRNA		8.7916	4.2526
RAI14	retinoic acid induced 14	Hs.15165	8.7584	4.0056
HSPA8	heat shock 70kDa protein 8	Hs.180414	8.7544	5.3559
ID3	inhibitor of DNA binding 3, dominant negative helix-loop-helix protein	Hs.76884	8.662	3.0338
PAK1IP1	PAK1 interacting protein 1	Hs.52256	8.6426	6.2093
HNRPA1	heterogeneous nuclear ribonucleoprotein A1		8.5958	4.1419
PRDX4	peroxiredoxin 4	Hs.83383	8.5721	4.1822
RPL38	ribosomal protein L38	Hs.425668	8.5404	3.8961
NBR2	neighbor of BRCA1 gene 2	Hs.321170	8.5097	7.8422
IMP-2	IGF-II mRNA-binding protein 2	Hs.30299	8.3643	6.3037
PUNC	putative neuronal cell adhesion molecule	Hs.189847	8.3588	4.7637
LAMR1	laminin receptor 1 (67kD, ribosomal protein SA)		8.348	5.3035

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
USP28	ubiquitin specific protease 28	Hs.142856	8.3187	4.1314
PRDX1	peroxiredoxin 1	Hs.180909	8.2866	3.8715
PIGPC1	p53-induced protein PIGPC1	Hs.303125	8.2262	4.7181
SFRP2	secreted frizzled-related protein 2	Hs.31386	8.186	6.1167
	Homo sapiens alpha-NAC gene for nascent polypeptide-associated complex component		8.1616	3.4958
PSIP1	PC4 and SFRS1 interacting protein 1	Hs.82110	8.1369	5.8686
MCM2	MCM2 minichromosome maintenance deficient 2, mitotin (S. cerevisiae)	Hs.57101	8.1332	6.9535
ERH	enhancer of rudimentary homolog (Drosophila)	Hs.433413	8.0448	3.5812
FLJ10377	hypothetical protein FLJ10377	Hs.274263	8.0292	3.585
C15orf15	chromosome 15 open reading frame 15	Hs.284162	8.0009	4.2542
RPL36AL	ribosomal protein L36a-like	Hs.419465	7.9765	3.1525
PHC1	polyhomeotic-like 1 (Drosophila)	Hs.305985	7.9451	3.1202
	Homo sapiens similar to laminin receptor 1 (67kD, ribosomal protein SA)		7.9431	5.8083
	Human DNA sequence from clone RP3-334F4 on chromosome 6		7.9329	5.8647
LRRN1	leucine rich repeat neuronal 1	Hs.126085	7.8855	6.1145
GRP58	glucose regulated protein, 58kDa	Hs.13751	7.8788	3.634
DSP	desmoplakin	Hs.349499	7.8722	4.792
CDC20	CDC20 cell division cycle 20 homolog (S. cerevisiae)	Hs.82906	7.8423	7.0917
HSPA4	heat shock 70kDa protein 4	Hs.90093	7.8213	5.0903
PSMD11	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11	Hs.90744	7.8084	3.1033
	Human DNA sequence from clone 522P13 on chromosome 6p21.31-22.3. Contains a 60S Ribosomal Protein L2.		7.7391	4.3755
KRT19	keratin 19	Hs.182265	7.7332	4.3622
TRIP13	thyroid hormone receptor interactor 13	Hs.6566	7.7075	3.2397
C20orf1	chromosome 20 open reading frame 1	Hs.9329	7.6512	4.8848
F2RL1	coagulation factor II (thrombin) receptor-like 1	Hs.154299	7.6243	4.5514
	Homo sapiens similar to 60S RIBOSOMAL PROTEIN L23A (H. sapiens) (LOC65880), mRNA		7.6189	3.9503
NOLA3	nucleolar protein family A, member 3 (H/ACA small nucleolar RNPs)	Hs.14317	7.5624	4.1481
	Human DNA sequence from clone RP11-352D3 on chromosome 20. Contains a 60S ribosomal protein L21 (RPL2).		7.539	3.7055
Jade-1	PHD protein Jade-1	Hs.238246	7.5164	4.3428
MRPL27	mitochondrial ribosomal protein L27	Hs.7736	7.505	3.6855
SEC61G	Sec61 gamma	Hs.9950	7.4484	4.0145
SNRPD3	small nuclear ribonucleoprotein D3 polypeptide 18kDa	Hs.1575	7.4432	3.6238
PCNP	PEST-containing nuclear protein	Hs.71618	7.4249	3.2268
PRO1855	hypothetical protein PRO1855	Hs.283558	7.4145	4.523
PHIP	pleckstrin homology domain interacting protein	Hs.10177	7.4053	4.5647
HMG1L10	high-mobility group (nonhistone chromosomal) protein 1-like 10		7.4048	4.2296
LASP1	LIM and SH3 protein 1	Hs.334851	7.3698	4.3823
GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	Hs.82285	7.3571	3.0676
ENO1	enolase 1, (alpha)	Hs.254105	7.3414	5.1802
ADAM19	a disintegrin and metalloproteinase domain 19 (meltrin beta)	Hs.278679	7.3083	4.2188
GA17	dendritic cell protein		7.2877	4.1859
MCM7	MCM7 minichromosome maintenance deficient 7 (S. cerevisiae)	Hs.77152	7.2792	5.6728

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
VDAC2	voltage-dependent anion channel 2	Hs.78902	7.2433	4.6981
KIAA0179	KIAA0179 protein	Hs.152629	7.21	4.1183
RPLP0	ribosomal protein, large, P0		7.1663	5.8115
MGST1	microsomal glutathione S-transferase 1	Hs.389700	7.158	4.587
TAF15	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa	Hs.381044	7.0972	3.2039
API5	apoptosis inhibitor 5	Hs.227913	7.0571	4.3763
RPL19	ribosomal protein L19	Hs.426977	7.0543	4.395
CCNA2	cyclin A2	Hs.85137	7.0351	3.3263
	Homo sapiens c33.42 unnamed HERV-H protein mRNA, partial cds		7.0218	20.4077
PSMB6	proteasome (prosome, macropain) subunit, beta type, 6	Hs.77060	7.0126	3.9363
RPA1	replication protein A1, 70kDa	Hs.84318	7.0075	4.8865
TK1	thymidine kinase 1, soluble	Hs.105097	6.9965	6.953
NPM3	nucleophosmin/nucleoplasmin, 3	Hs.90691	6.9806	4.0789
	Human DNA sequence from clone RP1-269M15 on chromosome 20q12-13.12. Contains a gene similar to peptid.		6.9789	5.0146
FLJ12484	hypothetical protein FLJ12484	Hs.37747	6.976	4.7357
HAND1	heart and neural crest derivatives expressed 1	Hs.152531	6.9567	3.178
IF2	translation initiation factor IF2	Hs.158688	6.9457	3.8838
TEBP	unactive progesterone receptor, 23 kD	Hs.278270	6.9419	3.1811
KIAA0020	minor histocompatibility antigen HA-8	Hs.2471	6.9309	3.7341
	Homo sapiens similar to ribosomal protein L23a (H. sapiens) (LOC65638), mRNA		6.9219	4.203
TNNT1	troponin T1, skeletal, slow	Hs.73980	6.8795	5.5485
PITX2	paired-like homeodomain transcription factor 2	Hs.92282	6.8032	3.0306
RPS23	ribosomal protein S23	Hs.3463	6.782	4.0483
	Homo sapiens putative human HLA class II associated protein I (PHAP1), mRNA		6.7527	3.7135
IFITM1	interferon induced transmembrane protein 1 (9-27)	Hs.458286	6.7212	7.3063
DKFZp762L031	hypothetical protein DKFZp762L0311	Hs.351623	6.6341	3.6789
PFN1	profilin 1	Hs.408943	6.6189	3.9552
PA2G4	proliferation-associated 2G4, 38kDa	Hs.343258	6.6106	3.6266
HNRPL	heterogeneous nuclear ribonucleoprotein L	Hs.2730	6.6088	3.2105
ACTA1	actin, alpha 1, skeletal muscle	Hs.1288	6.5897	6.8379
METAP1	methionyl aminopeptidase 1	Hs.82007	6.5638	4.303
PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase) 2	Hs.41270	6.5332	3.9661
	Homo sapiens similar to 60S RIBOSOMAL PROTEIN L23A (H. sapiens) (LOC65417), mRNA		6.5014	4.0788
SDCCAG8	serologically defined colon cancer antigen 8	Hs.300642	6.4882	3.5579
FLJ11046	hypothetical protein FLJ11046	Hs.16986	6.4805	3.0127
CKAP2	cytoskeleton associated protein 2	Hs.24641	6.4719	4.2204
MTHFD2	methylene tetrahydrofolate dehydrogenase (NAD+ dependent), methenyltetrahydrofolate cyclohydrolase	Hs.154672	6.4566	4.8834
PIF1	DNA helicase homolog PIF1	Hs.112160	6.4446	6.6647
MRPS7	mitochondrial ribosomal protein S7	Hs.71787	6.4113	3.5154
LYPLA1	lysophospholipase I	Hs.393360	6.3884	4.1219
GARS	glycyl-tRNA synthetase	Hs.293885	6.3879	3.9937
C7orf14	chromosome 7 open reading frame 14	Hs.84790	6.3741	4.1663
PRO1068	Hypothetical protein PRO1068		6.3429	3.31
MPB1	MYC promoter-binding protein 1		6.3354	6.4887

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
RPL4	ribosomal protein L4	Hs.286	6.3239	6.2255
Nanog	homeobox transcription factor Nanog	Hs.326290	6.3173	5.1194
MAK3P	likely ortholog of mouse Mak3p homolog (S. cerevisiae)	Hs.288932	6.3149	3.1316
RPS3A	ribosomal protein S3A	Hs.77039	6.3108	3.1347
CBX3	chromobox homolog 3 (HP1 gamma homolog, Drosophila)	Hs.406384	6.27	3.8758
	Human endogenous retrovirus pHE.1 (ERV9)		6.2605	4.4253
C21orf45	chromosome 21 open reading frame 45	Hs.49932	6.2391	4.0354
CKS1B	CDC28 protein kinase regulatory subunit 1B	Hs.348669	6.2083	3.5537
SET	SET translocation (myeloid leukemia-associated)	Hs.145279	6.1923	6.4237
FLJ10326	hypothetical protein FLJ10326	Hs.262823	6.1725	3.0784
LOC84549	RNA binding protein	Hs.77135	6.1689	3.5191
KPNB3	karyopherin (importin) beta 3	Hs.113503	6.1635	5.2102
ZNF117	zinc finger protein 117 (HPF9)	Hs.133011	6.123	4.5006
TEAD4	TEA domain family member 4	Hs.94865	6.1149	3.9091
HSPC117	hypothetical protein HSPC117	Hs.10729	6.089	3.5708
RAD21	RAD21 homolog (S. pombe)	Hs.81848	6.0875	3.3005
CCT3	chaperonin containing TCP1, subunit 3 (gamma)	Hs.1708	6.0834	5.028
HSPC163	HSPC163 protein	Hs.108854	6.0769	4.0512
RAMP	RA-regulated nuclear matrix-associated protein	Hs.126774	6.0676	4.5875
RPL10A	ribosomal protein L10a		6.0662	4.7655
FRSB	phenylalanyl-tRNA synthetase beta-subunit	Hs.9081	6.0412	3.546
SSRP1	structure specific recognition protein 1	Hs.79162	6.0172	3.0172
H2AFX	H2A histone family, member X	Hs.147097	6.0115	5.4197
ENTH	enthoprotin	Hs.132853	6.0027	4.2103
	Human DNA sequence from clone RP13-178D16 on chromosome X.		5.9915	3.0462
HMGB2	high-mobility group box 2	Hs.80684	5.9861	4.3673
HMGN2	high-mobility group nucleosomal binding domain 2	Hs.181163	5.9599	3.9082
	Human DNA sequence from clone RP11-16L21 on chromosome 9.		5.9588	3.3106
PTTG1	pituitary tumor-transforming 1	Hs.252587	5.958	5.3201
DRIL1	dead ringer-like 1 (Drosophila)	Hs.198515	5.9394	4.9808
FLJ20641	hypothetical protein FLJ20641	Hs.121553	5.9275	5.8264
FLJ21841	hypothetical protein FLJ21841	Hs.29076	5.9133	3.757
SNRPD2	small nuclear ribonucleoprotein D2 polypeptide 16.5kDa	Hs.424327	5.9112	3.7928
BRIX	BRIX	Hs.38114	5.9077	4.2369
PBEF	pre-B-cell colony-enhancing factor	Hs.239138	5.8609	3.3468
SERPINH1	serine (or cysteine) proteinase inhibitor, clade H (heat shock protein 47), member 1, (collagen binding protein 1)	Hs.9930	5.8493	5.7781
B3GNT7	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7	Hs.299329	5.8393	3.1599
PSMD14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14	Hs.178761	5.8212	4.1797
ANP32A	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	Hs.457988	5.7982	4.0879
GALNT1	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1)	Hs.80120	5.7882	4.6825
	Human DNA sequence from clone RP1-34P24 on chromosome 22.		5.7847	3.1726
FUSIP1	FUS interacting protein (serine-arginine rich) 1	Hs.3530	5.7592	3.3965
PSMB5	proteasome (prosome, macropain) subunit, beta type, 5		5.7573	4.5482
RPL24	ribosomal protein L24	Hs.184582	5.7409	4.6246

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
IMPDH2	IMP (inosine monophosphate) dehydrogenase 2	Hs.75432	5.7185	5.5777
RPL31	ribosomal protein L31	Hs.184014	5.6871	3.3864
	Human DNA sequence from clone RP11-102J14 on chromosome 20.		5.6831	3.3922
PSMA3	proteasome (prosome, macropain) subunit, alpha type, 3	Hs.346918	5.6703	5.5412
EEF1B2	eukaryotic translation elongation factor 1 beta 2	Hs.421608	5.6635	3.9278
	Human DNA sequence from clone 391O22 on chromosome 6p21.2-21.31.		5.6596	4.217
	Human DNA sequence from clone RP11-371L19 on chromosome 20.		5.6254	3.5281
RPL35A	ribosomal protein L35a	Hs.288544	5.6197	3.3447
HSPC111	hypothetical protein HSPC111	Hs.279918	5.5807	3.3189
COPB	coatomer protein complex, subunit beta	Hs.3059	5.5571	3.3839
MGC5627	hypothetical protein MGC5627	Hs.237971	5.554	5.3964
PPP2CA	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	Hs.91773	5.5484	3.4562
RRM2	ribonucleotide reductase M2 polypeptide	Hs.75319	5.5375	3.8071
	Sapiens thioredoxin delta 3 (TXN delta 3) mRNA, partial cds	Hs.395309	5.5326	3.1614
RPL9	ribosomal protein L9	Hs.426460	5.5307	3.1146
CRABP1	cellular retinoic acid binding protein 1	Hs.346950	5.4832	6.6444
CCNB2	cyclin B2	Hs.194698	5.4742	4.1123
BAF53A	BAF53	Hs.274350	5.4581	3.4022
MRPL32	mitochondrial ribosomal protein L32	Hs.50252	5.455	4.2648
ITGB5	integrin, beta 5	Hs.149846	5.4419	3.8255
	Homo sapiens thymosin beta-10 gene, 3′ end		5.4387	3.8365
EPRS	glutamyl-prolyl-tRNA synthetase	Hs.55921	5.428	5.7694
HNOEL-iso	HNOEL-iso protein	Hs.9315	5.4266	4.0303
RRM1	ribonucleotide reductase M1 polypeptide	Hs.2934	5.4238	5.5373
FLJ10006	hypothetical protein FLJ10006	Hs.5570	5.3752	3.5915
PGAM1	phosphoglycerate mutase 1 (brain)	Hs.457938	5.3577	3.9171
PFAS	phosphoribosylformylglycinamidine synthase (FGAR aminotransferase)	Hs.105478	5.3498	3.4139
FEN1	flap structure-specific endonuclease 1	Hs.4756	5.3438	6.2717
NMU	neuromedin U	Hs.418367	5.2972	3.925
	Human DNA sequence from clone RP3-406P24 on chromosome 6.		5.2836	3.2444
C20orf129	chromosome 20 open reading frame 129	Hs.70704	5.2795	5.2988
ATP1B3	ATPase, Na^+/K^+ transporting, beta 3 polypeptide	Hs.76941	5.2596	3.1735
HN1	hematological and neurological expressed 1	Hs.109706	5.2558	4.6087
DTYMK	deoxythymidylate kinase (thymidylate kinase)	Hs.79006	5.2432	3.8048
TDGF1	teratocarcinoma-derived growth factor 1	Hs.75561	5.2291	7.5328
CALU	calumenin	Hs.7753	5.2279	3.6995
HMBS	hydroxymethylbilane synthase	Hs.82609	5.2153	5.5948
TMSB10	thymosin, beta 10	Hs.76293	5.1879	3.9498
FBL	fibrillarin	Hs.99853	5.1798	4.6536
HDAC2	histone deacetylase 2	Hs.3352	5.1694	4.1544
RAB38	RAB38, member RAS oncogene family	Hs.108923	5.1376	3.3122
K-ALPHA-1	tubulin, alpha, ubiquitous	Hs.334842	5.1032	3.4594
IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	Hs.11223	5.0956	5.1678
PP	pyrophosphatase (inorganic)	Hs.184011	5.0907	3.0886
TUBB-5	beta 5-tubulin	Hs.179661	5.075	5.281
	Human DNA sequence from clone RP4-753D5 on chromosome 6p12.1-12.3.		5.0187	3.1236

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
CCT-5	chaperonin containing TCP1, subunit 5 (epsilon)	Hs.1600	5.0179	7.4194
MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	Hs.77171	4.999	4.8933
CDT1	DNA replication factor	Hs.122908	4.9676	3.1369
SLC39A1	solute carrier family 39 (zinc transporter), member 1	Hs.7854	4.9632	3.9787
SCYE1	small inducible cytokine subfamily E, member 1 (endothelial monocyte-activating)	Hs.333513	4.9498	3.8042
PIM1	pim-1 oncogene	Hs.81170	4.9063	3.4099
PMSCL1	polymyositis/scleroderma autoantigen 1, 75kDa	Hs.91728	4.9017	3.4002
ADSL	adenylosuccinate lyase	Hs.75527	4.8969	4.5402
CGI-94	comparative gene identification transcript 94	Hs.111449	4.8638	3.0047
HSSG1	Heat-shock suppressed protein 1		4.8632	4.3132
ZNF257	zinc finger protein 257	Hs.283900	4.8502	3.6248
EEF1A1	eukaryotic translation elongation factor 1 alpha 1	Hs.422118	4.8475	3.4583
FKBP3	FK506 binding protein 3, 25kDa	Hs.379557	4.8302	4.0163
SLC7A5	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	Hs.184601	4.8277	5.7672
UHRF1	ubiquitin-like, containing PHD and RING finger domains, 1	Hs.108106	4.827	3.0648
PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polypeptide	Hs.85701	4.8213	3.0922
RPS19	ribosomal protein S19	Hs.298262	4.809	4.4312
PAI-RBP1	PAI-1 mRNA-binding protein	Hs.165998	4.8059	3.8869
TACSTD1	tumor-associated calcium signal transducer 1	Hs.692	4.7731	3.6411
	Human DNA sequence from clone RP5-843L14 on chromosome 20.		4.6933	3.1038
PCNT1	pericentrin 1	Hs.184352	4.5864	3.2607
LOC51685	prothymosin a14	Hs.457743	4.5584	3.8078
MTHFD1	methylenetetrahydrofolate dehydrogenase (NADP+ dependent), methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	Hs.172665	4.5499	5.8678
	Human endogenous retrovirus mRNA for ORF		4.5458	4.4232
PSMB3	proteasome (prosome, macropain) subunit, beta type, 3	Hs.82793	4.5379	4.6985
MRPL12	mitochondrial ribosomal protein L12	Hs.109059	4.5373	5.2112
MAD2L2	MAD2 mitotic arrest deficient-like 2 (yeast)	Hs.19400	4.5047	6.3397
FUS	fusion, derived from t(12;16) malignant liposarcoma	Hs.99969	4.4953	3.1476
	Human DNA sequence from clone 551E13 on chromosome Xp11.2-11.3		4.4799	3.7248
RPS18	ribosomal protein S18	Hs.275865	4.436	4.8116
NIF3L1	NIF3 NGG1 interacting factor 3-like 1 (S. pombe)	Hs.21943	4.4351	3.293
MRPL17	mitochondrial ribosomal protein L17	Hs.10026	4.3239	3.6726
TRIP8	thyroid hormone receptor interactor 8	Hs.6685	4.3065	3.7629
FADS2	fatty acid desaturase 2	Hs.184641	4.2905	3.2499
CTBP2	C-terminal binding protein 2	Hs.171391	4.2707	3.1375
AGTRL1	angiotensin II receptor-like 1	Hs.9305	4.2638	3.7687
PDLIM1	PDZ and LIM domain 1 (elfin)	Hs.75807	4.2581	5.2093
UQCRH	ubiquinol-cytochrome c reductase hinge protein	Hs.73818	4.2482	3.1334
UCHL1	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	Hs.76118	4.2185	3.151
	Sapiens mRNA; cDNA DKFZp434O1317 (from clone DKFZp434O1317)	Hs.25362	4.2096	3.8012
	Homo sapiens prothymosin alpha (PTMA) gene, complete cds		4.2081	3.0005
	Human DNA sequence from clone RP5-1028D15 on chromosome 20.		4.2025	3.588
STK12	serine/threonine kinase 12	Hs.180655	4.1988	11.5657
RNASEH2A	ribonuclease H2, large subunit	Hs.25292	4.1979	4.9723
CENPH	centromere protein H	Hs.200395	4.1938	3.1486

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
SEMA3A	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	Hs.2414	4.1549	3.0951
TOMM40	translocase of outer mitochondrial membrane 40 homolog (yeast)	Hs.30928	4.1467	3.7005
	Homo sapiens chromosome 19, cosmid R31449		4.1036	3.6413
DDX36	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 36	Hs.9414	4.0607	3.3047
MSF	MLL septin-like fusion	Hs.181002	4.0546	3.1321
WDR12	WD repeat domain 12	Hs.73291	4.0318	4.2461
SLC1A5	solute carrier family 1 (neutral amino acid transporter), member 5	Hs.183556	4.0108	3.3858
CHC1	chromosome condensation 1	Hs.84746	3.9621	3.2663
RPL7	ribosomal protein L7	Hs.153	3.956	3.6098
UACA	uveal autoantigen with coiled-coil domains and ankyrin repeats	Hs.49753	3.913	3.338
	Human DNA sequence from clone 38C16 on chromosome 6q22.33-24.1.		3.8726	3.1904
MFAP2	microfibrillar-associated protein 2	Hs.389137	3.8318	4.87
DNAJB6	DnaJ (Hsp40) homolog, subfamily B, member 6	Hs.181195	3.8267	3.746
KIAA0117	KIAA0117 protein	Hs.322478	3.8165	3.3667
NXN	nucleoredoxin	Hs.374534	3.8065	3.2577
PSMA4	proteasome (prosome, macropain) subunit, alpha type, 4	Hs.251531	3.8052	3.7093
KIAA1553	KIAA1553 protein	Hs.184852	3.802	4.4838
ZNF43	zinc finger protein 43 (HTF6)	Hs.74107	3.7443	3.7323
COL18A1	collagen, type XVIII, alpha 1	Hs.78409	3.7102	3.741
RPS5	ribosomal protein S5	Hs.356019	3.6754	3.8684
CTSL2	cathepsin L2	Hs.87417	3.6367	4.304
VDP	vesicle docking protein p115	Hs.325948	3.6293	4.4329
MRPL37	mitochondrial ribosomal protein L37	Hs.4209	3.6128	3.3022
TUBB-4	tubulin beta-4	Hs.274398	3.5827	3.3015
ALPL	alkaline phosphatase, liver/bone/kidney	Hs.250769	3.5746	3.8268
SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	Hs.79748	3.5587	3.3374
GNA11	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	Hs.1686	3.5529	3.4664
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	Hs.151777	3.5242	3.9213
FLJ14502	TRAF4 associated factor 1	Hs.181466	3.5204	3.222
MGC14480	hypothetical protein MGC14480	Hs.37616	3.519	3.4166
SAD1	SnRNP assembly defective 1 homolog	Hs.12820	3.5134	3.3234
ABCF1	ATP-binding cassette, sub-family F (GCN20), member 1	Hs.9573	3.4952	3.1965
SMC6L1	SMC6 structural maintenance of chromosomes 6-like 1 (yeast)	Hs.34497	3.4339	3.0684
CGI-30	CGI-30 protein	Hs.406051	3.4226	3.466
OSBPL9	oxysterol binding protein-like 9	Hs.21938	3.3999	3.0572
FLJ23091	putative NFkB activating protein 373	Hs.250746	3.3894	3.6872
SC5DL	sterol-C5-desaturase (ERG3 delta-5-desaturase homolog, fungal)-like	Hs.288031	3.3769	3.1688
TACC3	transforming, acidic coiled-coil containing protein 3	Hs.104019	3.3682	3.1189
APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1	Hs.73722	3.3086	4.0841
CHORDC1	cysteine and histidine-rich domain (CHORD)-containing, zinc binding protein 1	Hs.22857	3.2837	3.4777
GDI2	GDP dissociation inhibitor 2	Hs.56845	3.2516	3.0181
LAPTM4B	lysosomal associated protein transmembrane 4 beta	Hs.296398	3.242	4.8001
NAV2	neuron navigator 2	Hs.23467	3.1751	3.6536
RPL29	ribosomal protein L29	Hs.430207	3.1732	3.2301
TKT	transketolase (Wernicke-Korsakoff syndrome)	Hs.89643	3.1717	3.9633

Table 3. Highly expressed genes in BG01 and BG02 cells (> threefold) (continued)

Gene	Description	UniGene	BG02	BG01
FDPS	farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, dimethylallyltranstransferase, geranyltranstransferase)	Hs.335918	3.1591	3.1378
EIF3S9	eukaryotic translation initiation factor 3, subunit 9 eta, 116kDa	Hs.57783	3.1329	3.374
HSPC242	hypothetical protein HSPC242	Hs.25199	3.1002	4.3694
DSG2	desmoglein 2	Hs.359784	3.0897	3.0869
HT007	uncharacterized hypothalamus protein HT007	Hs.24371	3.0866	3.6356
SLC5A6	solute carrier family 5 (sodium-dependent vitamin transporter), member 6	Hs.321579	3.0595	3.4608

Of the 96 transcripts examined, 31 transcripts were detected in both BG01 and BG02 and 65 transcripts were not detected in either line (Fig. 7B). Nodal, LeftyA, LeftyB, Cripto-1 (TDGF-1), Cerberus, activin receptor ALK-3, and SMAD (3, 5, and 6), all involved in the Nodal signaling pathway, were expressed in both BG01 and BG02. Several bone morphogenetic proteins (BMPs), BMP2, BMP4, and BMP9 were detected in undifferentiated ES cells; however, BMP1, BMP3, BMP5, BMP6, BMP7, BMP8, BMP10, and BMP11 were not expressed in BG01 or BG02. Basic helix-loop-helix transcription factor ID2, ID3, and ID4, but not ID1 were also found to be expressed in BG01 and BG02.

Gene Expression Profile by Human Stem Cell Array

A focused array with genes related to human stem cell populations was used to analyze gene expression in undifferentiated BG01 and BG02 cells. The array contains 266 known genes, including 86 that encode markers expressed by stem cells at various stages of differentiation, 96 growth factors and cytokines known to regulate stem cell growth, 36 genes encoding extracellular matrix molecules expressed at appropriate developmental stages, and genes encoding proteins such as cell cycle regulators that are thought to be involved in stem cell division. Positive controls and housekeeping genes were also included for normalization to generate relative expression profiles.

The results of the human stem cell array analysis are summarized in Figure 8. Of the 266 genes represented by the array, 102 genes were detected in BG01 and BG02. Genes highly expressed included 12 transcription factors, 7 cell cycle-related genes, 25 markers for stem and differentiated cells, 44 growth factors, adhesion molecules and

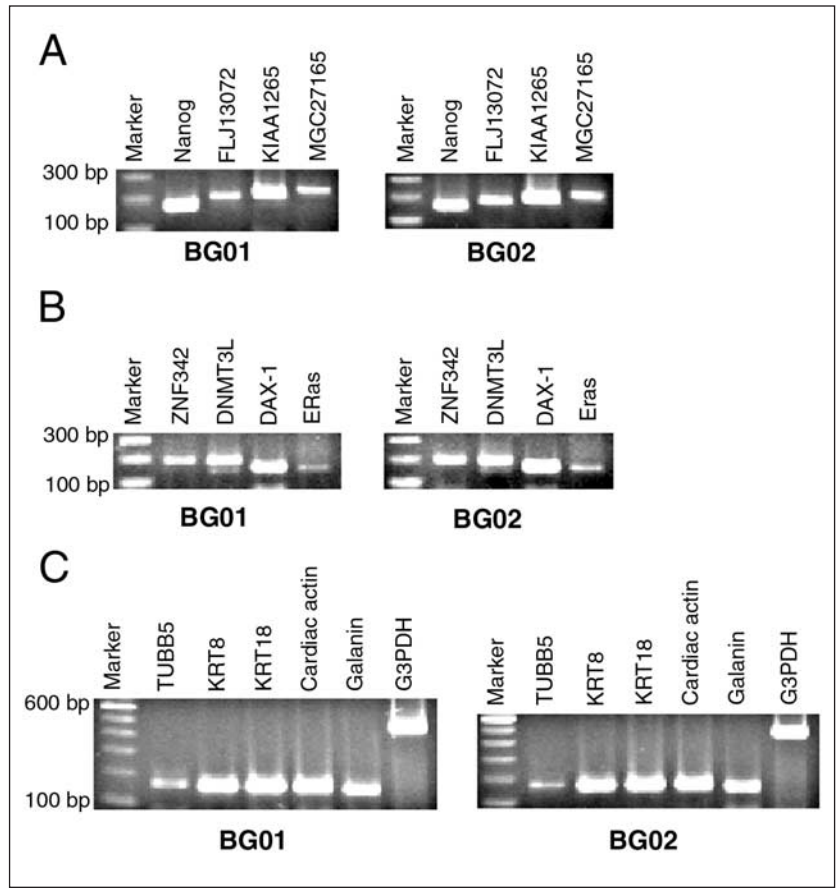

Figure 6. RT-PCR confirmation of representative genes detected by the large-scale microarray analysis of BG01 and BG02. *A) Nanog, a recently identified hypothetical protein FLJ12581, that is critical for maintaining pluripotency of mES cells and three other hypothetical proteins, FLJ13072, KIAA1265, and MGC27165, were expressed in both BG01 and BG02. B) ZNF342, DNMT3L, DAX-1, and Eras were also expressed. C) Expression of keratin 8, 14, and 19; cardiac actin; and tubulin (genes associated with differentiation) were also detected by RT-PCR.*

Table 4. Nineteen novel genes highly expressed (> threefold) identified by a large-scale microarray

Symbol	Genebank	Fold	Description
FLJ12519	AL122097	4	hypothetical protein FLJ12519 high ests in testes, cervix and tumors. Chr 2.
MGC5627	NM_024096	5	hypothetical protein MGC5627 predicted protein
FLJ20105	AK000112	4	hypothetical protein FLJ20105 X chr unknown function in B cell tumors
	AL133611	4	Homo sapiens mRNA; cDNA DKFZp434O1317 (from clone DKFZp434O1317) Chr11 has poly A in tumors and est no predicted function
FLJ20641	AK000648	6	hypothetical protein FLJ20641 Chr 12 hyothetical protein with helicase domains, highly enriched in EC libraries
AD024	NM_020675	8	AD024 protein unknown function full length
	AL031577	6	Human DNA sequence from clone RP3-391O22 on chromosome 6p21.2-21.31, complete sequence
HSPC163	NM_014184	6	HSPC163 protein INTEGRAL MEMBRANE PROTEIN (POTENTIAL). [TISSUE SPECIFICITY] EXPRESSED IN OOCYTES and adult brain member of the cornichon family (37% homology)
FLJ12505	NM_024749	23	hypothetical protein FLJ12505 Chr1 290 aa found in ES est library
	AL356414	7	Human DNA sequence from clone RP11-352D3 on chromosome 20 contains an 0S ribosomal protein L21
FLJ10156	AK001018	14	hypothetical protein FLJ10156 Chr 9 enriched in neuroblastomas
DKFZP586L0724	AL110271	17	DKFZP586L0724 protein Chr 7
HT007	NM_018480	3	uncharacterized hypothalamus protein HT007
M55914		6	MYC promoter binding protein 1
FLJ12684	NM_024534	3	hypothetical protein FLJ12684 Chr 4563aa very weakly similar to env protein
CGI-94	NM_016037	5	comparative gene identification transcript 94
HSPC117	AL050255	6	hypothetical protein HSPC117 Chr22 505 aa a family of uncharacterized proteins
HSPC111	NM_016391	5	hypothetical protein HSPC111 Chr 5 178aa conserved in c-elegans
FLJ21841	NM_024609	6	hypothetical protein FLJ21841 63% homology to nestin

cytokines, and 14 others. In particular, telomerase (TERT), telomerase-associated protein 1, sonic hedgehog, five members of the TGF-β superfamily and receptors, seven members of the FGF family and receptors, 3 WNT genes, and ESG1, an expressed sequence tag that is weakly similar to embryonic stem cell-specific gene 1, were highly expressed. Members of the FGF and leukemia inhibitory factor (LIF) families and their receptors (LIFR) that were expressed are listed in Table 5. Expression of LIF and LIFR was below the limit of detection, and expression of gp130 was very low. The spots for these transcripts, which can be identified in the array [27], are not visible in Figure 8.

Discussion

Despite the many potential uses of hES cell lines, only a limited number of lines are available and the properties of only a few of these have been described in the literature [5]. These include the lines available from Wicell (http://www.wicell.org/index.jsp) with more limited data available on lines from ESI (http://www.escellinternational.com) and from the Israel Institute of Technology (http://www.technion.ac.il). Of the four hES cell lines BG01, BG02, BG03, and BG04, derived by BresaGen from discarded embryos [6], two of them, BG01 and BG02, are currently available for academic research. In the present study, we report on characterization of these two hES cell lines. Both BG01 and BG02 express markers for undifferentiated ES cells, similar to other hES cell lines that have previously been described. BG01 and BG02 appear to be virtually identical in their ability to differentiate into ectodermal, mesodermal, and endodermal derivatives in vitro and in expression of cell surface antigens and transcription factors. Using a large-scale oligonucleotide microarray and small focused microarrays, we have confirmed expression of common markers and identified numerous additional genes that are expressed in undifferentiated BG01 and BG02 cells.

The potential of hES cells to contribute to the germline of chimeric organisms cannot be tested in humans and is not readily tested in nonhuman primates [2-4], and thus additional criteria for evaluating human and primate ES cells need to be used [4]. Human and subhuman primate ES cells have been defined as cells that are derived from pre-implantation of peri-implantation embryos, and can be maintained in culture for prolonged periods in an undifferentiated state, while retaining the capacity to differentiate into cells of all three embryonic germ layers [4]. Our results show that like the hES cell lines, BG01 and BG02 appear

A Image profile BG01 versus BG02

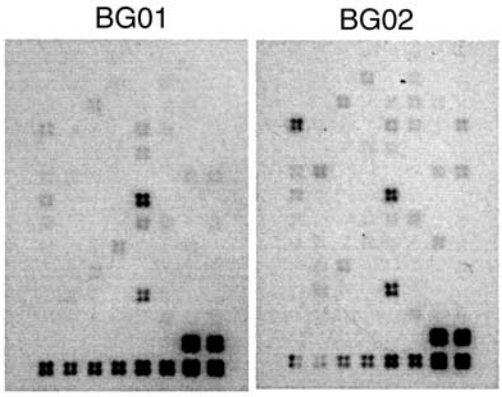

Figure 7. Gene expression profile of BG01 and BG02 by a focused human stem cell microarray. *A) Images of arrays hybridized using BG01 and BG02 RNA. B) Summary of genes expressed by BG01 and BG02.*

B Genes detected by TGF-β/BMP array (31 genes)

Category	Genes
TGF-β superfamily cytokine (8)	Nodal, LeftyA, LEftyB, INHA, BMP2, BMP4, BMP9, GDF3
Molecules regulating signaling of TGF-β superfamily (4)	Cerberus, Cripto-1 (TDGF-1), FST (Follistatin), NMA
SMAD (4)	Smad3, Smad4 (DPC4), Samd5, Samd6
Smad target genes (13)	CDC25A, p21Waf1(p21Cip1), COL1A2, COL3A1, FOS, JunB, Myc, TGIF, TIMP1, ID2, ID3, ID4, SOX4
Receptor (2)	ALK-3, BMPR2

morphologically and antigenically similar to the previously described hES cell lines. Like other hES cell lines [26] and unlike mouse ES cells, BG01 and BG02 cells are larger, grow more slowly, and grow in fewer layers. BG01 and BG02 express markers of undifferentiated ES cells like Sox-2, Oct-3/4, Nanog, TERT, SSEA-3, and SSEA-4; lack expression of markers of differentiation; can be maintained in culture for over 40 passages; and retain the ability to differentiate and express markers characteristic of ectoderm (TuJ1), endoderm (AFP), and mesoderm (cTnI). Like other cell lines tested [26, 28], BG01 and BG02 express galanin, Keratin 8 and 18, and several additional novel markers of the ES state such as Nanog and ZFN342. Similar to other ES lines, BG01 and BG02 can be cryopreserved and propagated extensively, and preliminary results suggest that clonal derivatives can be isolated. Like other ES cells that have been described, BG01 and BG02 do not appear to require LIF for their propagation and survival, and expression of LIFR or gp130 is low or absent (Fig. 7; Table 5; data not shown).

Thus BG01 and BG02, like the limited number of other hES lines, fulfill the minimal definition of an ES cell. It is important to note that although the NIH hES cell registry includes 78 derivations (http://www.stemcells.nih.gov/registry/eligibilityCriteria.asp), most have not been demonstrated to meet the minimal definition of stem cells. A subset of the derivations have, however, been described as hES cell lines in the literature, primarily those from Wicell and ESI, that can be cryopreserved and maintained in culture, undifferentiated, for several months.

Direct comparison with other lines and publicly available databases suggest that BG01 and BG02 have properties that are very similar to other hES cells [4, 26]. However, BG01 and BG02 cells in vitro require a much higher density of MEFs in the feeder layer in order to maintain in an undifferentiated state. In contrast to other hES cell lines, BG01 and BG02 can be passaged by the use of trypsin, which works as well as nonenzymatic cell dissociation buffers. More importantly, BG01 and BG02 recovered more rapidly from frozen vials (within 2-3 days) as compared with GE01 and GE04 cell lines. In general, these cells are easy to grow and maintain in vitro, which could be an advantage of these lines as compared with other hES cells that require mechanical dissociation or special handling.

Gene expression patterns for the BG01 and BG02 lines were analyzed by immunocytochemistry, RT-PCR, and two separate microarray platforms. Three hundred seventy-three genes were identified as being highly expressed in both BG01 and BG02. Included were 92 stemness genes known to be expressed in mES or hES cells such as OCT-3/4, NANOG, Cripto/TDGF1, Cx43, and galanin [28]. The same genes were also overexpressed in an additional four hES lines (TE06, GE01, GE09, and a pooled set of subclones derived from GE01, GE09, and GE07). Differentially expressed genes include Nanog (a recently identified protein FLJ12581 critical for maintaining pluripotency of mouse ES cells) and other hypothetical

A Image profile BG01

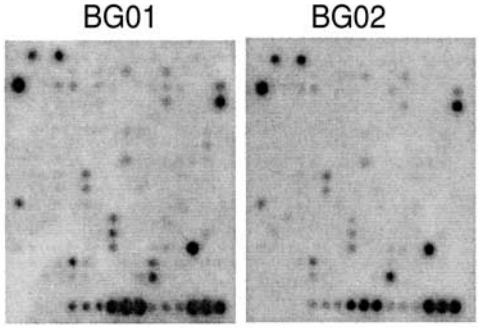

Figure 8. Gene expression profile of BG01 and BG02 by a focused TGF-β/BMP microarray. *A) Images of arrays hybridized using BG01 and BG02 RNA. B) Summary of genes expressed by BG01 and BG02.*

B Genes detected by human stem cell array (102 genes)

Category	Genes
Transcription factors (12)	FOXH1, SOX17, SOX18, SOX3, SOX6 SOX10, SOX2, NKX2-5, NKX2B, PAX6 POU3F1, PROX1
Cell cycle (7)	CDK4, CDKN1B, CDKN2A, CXCL12, CXCR4, MDM2, RBL2
Molecular markers for stem and differentiated cells (25)	ACTA2, ACTB, ACTG2, CD24, CD34, CD9, CNP, EGFR, EGR2, FGFR4, GFAP, IL6CT, KRT14, KRT17, KRT8, MAP2, NCAM2, PTPRC, SLC1A6, SLC2AI, SNAI1, SNAI2, TEP1, TERT, THY1
Growth factors, adhesion molecules and cytokines (44)	ACVR2, BMP1, BMP3, BMP5, BMPR1B, CNTF, CNTFR, FGF18, FGF19, FGF20, FGF9, FGFR2, FGFR3, FLJ10314 (Nodal), FZD2, GDF1, GDF2, IGF1R, IGF2, IGF2R, INS, LOC145957 (Neuregulin), NEUROG1, NRG1, PTCH2, TGFB3, TGFBR3, VEGF, WNT5A, WNT7B, WNT8A, CDH2, CDH4, CDH5, CTNNA1, CTNND2, ITGA5, ITGA6, ITGAV, ITGB1, ITGB5, NTRK3, PECAM1, VCAM1
Others (14)	C3orf4, CCNE1, CCNE2, , DNMT2, DNMT3L, ESG1, GAPD, HSPA9B, NOG, PPIA, PUM1, RPL13A, SHH, TINF2

proteins, KIAA1573 and MGC27165, which were all highly expressed in both BG01 and BG02. Both cell lines also expressed ZNF342, DNMT3L, and DAX-1, which were also confirmed by RT-PCR. Genes highly expressed in BG01 and BG02 also included numerous cell signaling/cell cycle/cell development-related genes, metabolism genes including DNA replication and DNA repair enzymes, RNA-related ribosomal genes, metabolic activity-related genes, transcription factors, and immune response genes. Also present were 19 novel genes with unknown function, illustrating the value of a large-scale gene expression analysis and the potential for identifying novel pathways of regulation.

Among the genes that were differentially expressed between BG01 and BG02 are members of the TGF-β superfamily member Nodal and its antagonist Lefty. Nodal signaling plays important roles in early embryonic development, patterning, and left-right axis positioning, as well as in the early stages of ES cell development [29-32]. Expression of genes related to the Nodal signaling pathway in undifferentiated BG01 and BG02 cells was studied with a focused microarray containing probes for TGF-β superfamily members and key proteins involved in the TGF-β signal-transduction pathway. Nodal; LeftyA; LeftyB; Cripto-1 (TDGF-1); Cerberus; activin receptors ALK-3; and SMAD3, SMAD5 and SMAF6, all molecules of the Nodal signaling pathway [33, 34], were expressed in both BG01 and BG02. While all components are present, this pathway is actively inhibited by negative regulators such as TDGF-1 and Nodal. Surprisingly, noggin does not appear to play as important a role in hES cell differentiation as has been proposed in rodent and xenopus studies [35-37]. Overall, the expression patterns for the TGF-β superfamily were virtually identical in BG01 and BG02, further reinforcing the similarity of the two lines.

Human ES cells are likely to be an important resource for biomedical research over the next decade, since these cells will allow studies of differentiation, gene expression, and biochemical pathways to be performed using human material. Cells that have been differentiated from hES cells are also likely to be useful for a wide range of mundane but valuable purposes such as screening drugs and antiviral agents. Thus, in addition to the therapeutic possibilities, there are important potential uses of hES cells for basic research. Although it is potentially possible to generate a large number of hES cell lines, the number of lines that

Table 5. BG01 and BG02 cells express several FGFs and their receptors but not LIF and its receptor LIFR using a human stem cell array

Gene	Description	Expressed
FGF9	Fibroblast growth factor 9	+
FGF18	Fibroblast growth factor 18	+
FGF19	Fibroblast growth factor 19	+
FGF20	Fibroblast growth factor 20	+
FGFR2	Fibroblast growth factor receptor 2	+
FGFR3	Fibroblast growth factor receptor 3	+
IL-6ST	Interleukin 6 signal transducer (gp130)	+
FGF1	Fibroblast growth factor 1 (acidic)	—
FGF2	Fibroblast growth factor 2 (basic)	—
FGF3	Fibroblast growth factor 3	—
FGF4	Fibroblast growth factor 4	—
FGF5	Fibroblast growth factor 5	—
FGF6	Fibroblast growth factor 6	—
FGF7	Fibroblast growth factor 7	—
FGF8	Fibroblast growth factor 8	—
FGF10	Fibroblast growth factor 10	—
FGF11	Fibroblast growth factor 11	—
FGF12	Fibroblast growth factor 12	—
FGF14	Fibroblast growth factor 14	—
FGF16	Fibroblast growth factor 16	—
FGF17	Fibroblast growth factor 17	—
FGF21	Fibroblast growth factor 21	—
FGF23	Fibroblast growth factor 23	—
FGFR1	Fibroblast growth factor receptor 1	±
FGFR4	Fibroblast growth factor receptor 4	±
LIF	Leukemia inhibitory factor	—
LIFR	Leukemia inhibitory factor (LIFR)	—

will be needed to be representative is unclear. Human ES cell lines may differ from each other due to differences in conditions and developmental stage under which they were isolated, in addition to male-female differences and differences related to genetic variations. While it is expected that the similarities among hES cell lines would be much greater than their differences, even small differences between hES cell lines may be critical in determining their utility. Although the NIH registry contains 78 putative stem cell lines(http://stemcells.nih.gov/registry/eligibilityCriteria.asp), very few of these lines have been described in the literature.

Our results show that BG01 and BG02 are generally similar to the other hES cell lines for which data on characterization are available. Both cell lines can be easily maintained in an undifferentiated state, grow rapidly, and readily differentiate into all major phenotypes, suggesting that these lines can be added to the short list of validated, potentially useful hES cell lines.

Acknowledgments

We thank members of our laboratories for constant stimulating discussions. We acknowledge the support of NIH grant R24DK063689 to BresaGen Inc.

References

1 Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981;292:154-156.

2 Thomson JA, Kalishman J, Golos TG et al. Isolation of a primate embryonic stem cell line. Proc Natl Acad Sci USA 1995;92:7844-7848.

3 Thomson JA, Kalishman J, Golos TG et al. Pluripotent cell lines derived from common marmoset (Callithrix jacchus) blastocysts. Biol Reprod 1996;55:254-259.

4 Thomson JA, Itskovitz-Eldor J, Shapiro SS et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1140-1147.

5 Carpenter MK, Rosler E, Rao MS. Characterization and differentiation of human embryonic stem cells. Cloning Stem Cells 2003;5:79-88.

6 Mitalipova M, Calhoun J, Shin S et al. Human embryonic stem cell lines derived from discarded embryos. Stem Cells 2003;21:521-526.

7 Amit M, Itskovitz-Eldor J. Derivation and spontaneous differentiation of human embryonic stem cells. J Anat 2002;200:225-232.

8 Carpenter MK, Rosler E, Fisk GJ et al. Properties of four human ES cell lines maintained in a feeder free culture system. Dev Dyn 2004;229:243-258.

9 Rosler E, Fisk G, Ares X et al. Long term culture of human embryonic stem cells in feeder-free conditions. Dev Dyn 2004;229:259-274.

10 Tanaka TS, Kunath T, Kimber WL et al. Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. Genome Res 2002;12:1921-1928.

11 Xu C, Inokuma MS, Denham J et al. Feeder-free growth of undifferentiated human embryonic stem cells. Nat Biotechnol 2001;19:971-974.

12 Eiges R, Schuldiner M, Drukker M et al. Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. Curr Biol 2001;11:514-518.

13 Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol 2003;21:319-321.

14 Carpenter MK, Inokuma MS, Denham J et al. Enrichment of neurons and neural precursors from human embryonic stem cells. Exp Neurol 2001;172:383-397.

15 Kaufman DS, Hanson ET, Lewis RL et al. Hematopoietic colony-forming cells derived from human embryonic stem cells. Proc Natl Acad Sci USA 2001;98:10716-10721.

16 Kehat I, Kenyagin-Karsenti D, Snir M et al. Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. J Clin Invest 2001;108:407-414.

17 Odorico JS, Kaufman DS, Thomson JA. Multilineage differentiation from human embryonic stem cell lines. Stem Cells 2001;19:193-204.

18 Rambhatla L, Chiu CP, Kundu P et al. Generation of hepatocyte-like cells from human embryonic stem cells. Cell Transplant 2003;12:1-11.

19 Reubinoff BE, Itsykson P, Turetsky T et al. Neural progenitors from human embryonic stem cells. Nat Biotechnol 2001;19:1134-1140.

20 Xu C, Police S, Rao N et al. Characterization and enrichment of cardiomyocytes derived from human embryonic stem cells. Circ Res 2002;91:501-508.

21 Zhang SC, Wernig M, Duncan ID et al. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol 2001;19:1129-1133.

22 Kumar A, Novoselov V, Celeste AJ et al. Nodal signaling uses activin and transforming growth factor-beta receptor-regulated Smads. J Biol Chem 2001;276:656-661.

23 McWhir J, Schnieke AE, Ansell R et al. Selective ablation of differentiated cells permits isolation of embryonic stem cell lines from murine embryos with a non-permissive genetic background. Nat Genet 1996;14:223-226.

24 Vassilieva S, Guan K, Pich U et al. Establishment of SSEA-1- and Oct-4-expressing rat embryonic stem-like cell lines and effects of cytokines of the IL-6 family on clonal growth. Exp Cell Res 2000;258:361-373.

25 Zeng X, Chen J, Sanchez et al. Stable expression of hrGFP by mouse embryonic stem cells: promoter activity in the undifferentiated state and during dopaminergic neural differentiation. STEM CELLS 2003;21:647-653.

26 Ginis I, Luo Y, Miura T et al. Differences between human and mouse embryonic stem cells. Dev Biol 2004 (in press).

27 Luo Y, Cai J, Ginis I et al. Designing, testing and validating a stem cell microarray for characterization of neural stem cells and progenitor cells. STEM CELLS 2003;21:575-587.

28 Bhattacharya B, Miura T, Brandenberger R et al. Gene expression in human embryonic stem cell lines: unique molecular signature. Blood 2004;103:2956-2964.

29 Juan H, Hamada H. Roles of nodal-lefty regulatory loops in embryonic patterning of vertebrates. Genes Cells 2001;6:923-930.

30 Schier AF, Shen MM. Nodal signalling in vertebrate development. Nature 2000;403:385-389.

31 Tiedemann H, Asashima M, Grunz H et al. Pluripotent cells (stem cells) and their determination and differentiation in early vertebrate embryogenesis. Dev Growth Differ 2001;43:469-502.

32 Gritsman K, Talbot WS, Schier AF. Nodal signaling patterns the organizer. Development 2000;127:921-932.

33 Kumar A, Novoselov V, Celeste AJ et al. Nodal signaling uses activin and transforming growth factor-beta receptor-regulated Smads. J Biol Chem 2001;276:656-661.

34 Gray PC, Harrison CA, Vale W. Cripto forms a complex with activin and type II activin receptors and can block activin signaling. Proc Natl Acad Sci USA 2003;100:5193-5198.

35 Wessely O, De Robertis EM. Neural plate patterning by secreted signals. Neuron 2002;14;33:489-491.

36 Giudice G. Conserved cellular and molecular mechanisms in development. Cell Biol Int 2001;25:1081-1090.

37 Sharpe C. Noggin—the neural inducer or a modifier of neural induction? Bioessays 1994;16:159-160.