A Quest for Human and Mouse Embryonic Stem Cell-specific Proteins*

Dennis Van Hoof‡§, Robert Passier‡¶, Dorien Ward-Van Oostwaard‡¶, Martijn W. H. Pinkse§, Albert J. R. Heck§, Christine L. Mummery‡¶, and Jeroen Krijgsveeld§**

Embryonic stem cells (ESCs) are of immense interest as they can proliferate indefinitely in vitro and give rise to any adult cell type, serving as a potentially unlimited source for tissue replacement in regenerative medicine. Extensive analyses of numerous human and mouse ESC lines have shown generic similarities and differences at both the transcriptional and functional level. However, comprehensive proteome analyses are missing or are restricted to mouse ESCs. Here we have used an extensive proteomic approach to search for ESC-specific proteins by analyzing the differential protein expression profiles of human and mouse ESCs and their differentiated derivatives. The data sets comprise 1,775 non-redundant proteins identified in human ESCs, 1,532 in differentiated human ESCs, 1,871 in mouse ESCs, and 1,552 in differentiated mouse ESCs with a false positive rate of <0.2%. Comparison of the data sets distinguished 191 proteins exclusively identified in both human and mouse ESCs but not in their differentiated derivatives. Besides well known ESC benchmarks, this subset included many uncharacterized proteins, some of which may be novel ESC-specific markers. To complement the mass spectrometric approach, differential expression of a selection of these proteins was confirmed by Western blotting, immunofluorescence confocal microscopy, and fluorescence-activated cell sorting. Additionally two other independently isolated and cultured human ESC lines as well as their differentiated derivatives were monitored for differential expression of selected proteins. Some of these proteins were identified exclusively in ESCs of all three human lines and may thus serve as generic ESC markers. Our wide scale proteomic approach enabled us to screen thousands of proteins rapidly and select putative ESC-associated proteins for further analysis. Validation by three independent conventional protein analysis techniques shows that our methodology is robust, provides an excellent tool to characterize ESCs at the protein level, and may disclose novel ESC-specific benchmarks. Molecular & Cellular Proteomics 5:1261–1273, 2006.

Embryonic stem cells (ESCs)1 (1–5) are characterized by their ability to self-renew indefinitely and differentiate into somatic (6) and germ cells (7) of the adult individual. ESCs are derived from the inner cell mass of blastocyst stage embryos and are generally co-cultured on a monolayer of mitotically inactivated mouse embryonic fibroblast feeder cells to inhibit spontaneous differentiation. The first ESC lines were derived from mice (1, 2) and have since been subject to extensive research, serving as a model system to study differentiation in early mammalian development and a vehicle to modify the germ line in mice. Differences between pluripotent human and mouse embryonal carcinoma cells derived from malignant teratocarcinomas predicted that there would be differences between human and mouse ESCs (hESCs and mESCs, respectively). Initial characterization following their first derivation from human blastocysts (3–5) confirmed this was the case despite growth and co-culture with feeder cells supporting indefinite proliferation for hESCs as it had for mESCs.

Signaling pathways mediating self-renewal and differentiation to some cell lineages have been identified in mESCs (8), but their activity is only partially conserved in hESCs (9, 10). For instance, in contrast to hESCs, mESCs retain their undifferentiated state in the absence of feeder cells when cultured in serum-containing medium supplemented with leukemia inhibitory factor (LIF) (8, 11) via the signal transducer and activator of transcription-3 pathway. Bone morphogenetic proteins in combination with LIF can replace the requirement for

1 The abbreviations used are: ESC, embryonic stem cell; Dif-ESC, differentiated embryonic stem cell; Dif-hESC, differentiated human embryonic stem cell; Dif-mESC, differentiated mouse embryonic stem cell; FACS, fluorescence-activated cell sorting; hESC, human embryonic stem cell; LIF, leukemia inhibitory factor; mESC, mouse embryonic stem cell; DMEM, Dulbecco’s modified Eagle’s medium; HRP, horseradish peroxidase; IPI, International Protein Index; MCM, minichromosome maintenance; ACTB, β-actin; TOP2A, topoisomerase IIα; CACYBP, calcyclin-binding protein; MSH6, mutS homolog 6; RPS4X, X-linked ribosomal protein S4; KPN2A, karoyphelin-α2; RANBP2, Ran-binding protein 2; MTHFD1, methylenetetrahydrofolate dehydrogenase 1.
feeder cells and serum entirely for mESCs (12) but not for hESCs. On the other hand, hESCs have the ability to form trophoblast cells in response to bone morphogenetic proteins, whereas mESCs do not (for a review, see Ref. 13). Furthermore hESCs and mESCs differ in the expression of several cell surface antigens. For instance, stage-specific embryonic antigen-1 is expressed by undifferentiated mESCs and is down-regulated upon differentiation but is not expressed by hESCs until after differentiation. Conversely stage-specific embryonic antigen-3 and -4, TRA-1-60, TRA-1-81, and GCTM2 are examples of cell surface antigens specifically expressed by hESCs but not by mESCs or differentiated human cells.

Despite the differences between human and mouse ESCs, they share the ability to form many, if not all, somatic cell types. Furthermore they both express multiple transcription factors associated with this pluripotentiality, including Oct4 (POU5F1) (14), Nanog (15, 16), and Sox2 (17, 18), which collaborate in the regulation of pluripotency and self-renewal (19). In addition, high telomerase activity indicates an unlimited proliferative capacity of both hESCs and mESCs in vitro.

The combination of an unlimited ability to expand in culture and form derivatives of the three primary germ layers that give rise to all cells of the adult individual makes hESCs a particularly attractive source of cells for regenerative medicine (20). Therefore, detailed characterization of ESCs and their derivatives is vital to gain insight into the molecular processes underlying maintenance of the undifferentiated state of ESCs (for in vitro expansion) and cell lineage specification (for in vitro tissue generation).

To date, the most comprehensive comparative analyses of human and mouse ESC lines have been limited to a generic transcriptional level (21–23). We set out to complement these microarray studies at the translational level. Here we display the most prevalent proteins in the proteomes of undifferentiated human and mouse ESCs and their early differentiated derivatives using FT-ICR-MS/MS. In this study we extend earlier reports on the proteomic analysis of mESCs (24, 25) by monitoring the changes in protein expression upon differentiation and by comparing human and mouse cells during this process. Our study provides the largest and most accurate proteome data sets of hESCs and mESCs to date. From these data sets, proteins emerge that had not yet been typically associated with undifferentiated ESCs before. Moreover for a selection of proteins we have confirmed their differential expression in ESCs and differentiated ESCs (Dif-ESCs) by three independent means, i.e. Western blotting, immunofluorescence confocal microscopy, and fluorescence-activated cell sorting (FACS) (Fig. 1). Benchmarking ESCs at the protein level before or during differentiation is essential for characterizing existing and new hESC lines. Furthermore comparison of human and mouse ESCs may identify fundamental differences that might explain their distinct and common behavior. In addition, it provides a protein signature for individual ESC lines to monitor their stability in long term culture, uniqueness, and epigenetic status (26, 27).

**EXPERIMENTAL PROCEDURES**

**Materials**—DMEM, knock-out serum replacement, fetal calf serum, Glutamax, non-essential amino acids, β-mercaptoethanol, penicillin/streptomycin, human LIF, basic fibroblast growth factor, L-glutamine, insulin-transferrin-selenium solution (Invitrogen), Plasmanate (albumin fraction) (Bayer USA), Hyclone fetal calf serum (Perbio), mitomycin-C (Sigma), all fluorescently labeled antibodies and HRP-coupled rabbit anti-goat antibody (Jackson Immunoresearch), HRP-coupled goat anti-mouse (BD Biosciences), and HRP-coupled donkey anti-rabbit (Amersham Biosciences) were purchased from commercial resources. Cold water fish gelatin (Sigma) was kindly provided by K. W. Rodenburg. The primary antibodies were obtained from manufacturers or were generous gifts from the authors as listed in Table I.

**Cell Culture**—HES-2 hESCs were cultured in DMEM, 20% Hyclone...
fetal calf serum, L-glutamine, non-essential amino acids, insulin-transferrin-selenium, 50 nM β-mercaptoethanol, penicillin/streptomycin and passed mechanically by cut-and-paste on a monolayer of irradiated or mitomycin-C-treated mouse embryonic fibroblast feeder cells as described previously (4, 45). Differentiation was induced by culture for 12 days in the absence of feeder cells. D3 mESCs were mitotically inactivated human foreskin fibroblast cells, and passaged as HES-2 hESCs to passage 14, then transferred to a monolayer of knock-out DMEM, 20% knock-out serum replacement, 10 ng/ml basic fibroblast growth factor, Glutamax, non-essential amino acids, 50 nM β-mercaptoethanol, penicillin/streptomycin. Differentiation was induced by 5-day suspension culture after which the cell aggregates were seeded on gelatin-coated plates and grown for an additional 7 days.

**Preparation of Cell Extracts**—The centers of HES-2 and NL-HESC-01 hESC colonies, containing relatively high numbers of differentiated cells, were manually excised and discarded. Undifferentiated cells from all hESC lines were dissected from the monolayers of feeder cells by precise excision and collected in PBS on ice. Undifferentiated mESCs as well as differentiated human and mouse ESCs were harvested in PBS on ice. The cells were lysed by sonication for 10 s in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 50 mM NaF, 0.5% Nonidet P-40, 1 mM NaVO₃, 1 mM DTT, 1 mM PMSF, Roche Diagnostics protease inhibitor mixture) and subsequent incubation on ice for 30 min.

**In-gel Digestion and Peptide Sample Preparation**—Approximately 1.5 mg of protein of each sample was subjected to SDS-PAGE (47) using 12.5% polyacrylamide gels. Gels were lightly stained with Coomassie Blue, and each lane was cut into 26 parts. Each of these was cut down to smaller pieces and dehydrated with acetonitrile. Proteins were digested in-gel overnight at 37 °C using sequencing grade bovine trypsin (Roche Diagnostics; 10 μg/ml in 50 mM ammonium bicarbonate, pH 8.5). The supernatant was used for analysis by nanoflow liquid chromatography and FT-ICR-MS/MS.

**On-line Nanoflow Liquid Chromatography FT-ICR-MS/MS and Protein Identification**—Peptides generated by in-gel digestion were analyzed by nanoflow liquid chromatography using an Agilent 1100 HPLC...
system (Agilent Technologies) coupled on line to a 7-tesla LTQ-FT mass spectrometer (Thermo Electron, Bremen, Germany). The system was operated in a set-up essentially as described previously (48). ReproSil-Pur C18-AQ, 3 μm (Dr. Maisch GmbH, Ammerbuch, Germany) was used as a resin for capillary reversed phase chromatography. Peptides were trapped at 5 μl/min on a 1-cm column (100-μm internal diameter, packed in house) and eluted to a 15-cm column (50-μm internal diameter, packed in house) at ~150 nl/min in a 60-min gradient from 0 to 50% acetonitrile in 0.1 M acetic acid. The eluent was sprayed via emitter tips (made in house) butt-connected to the analytical column. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition. Full scan MS spectra were acquired in FT-ICR with a resolution of 20,000 at a target value of 2,000,000. The three most intense ions were then isolated for accurate mass measurements by a selected ion monitoring scan in FT-ICR with a resolution of 50,000 at a target accumulation value of 50,000. These ions were then fragmented in the linear ion trap using collision-induced dissociation at a target value of 15,000. In a postanalysis process, raw data were converted to peak lists using BioworksBrowser software, Version 3.1. For protein identification, MS/MS data were submitted to the International Protein Index (IPI) human (release 2.36, 36,972 entries, October 2004) and IPI mouse databases (release 1.29, 41,676 entries, October 2004) using Mascot Version 2.1 (Matrix Science) with the following settings: 1–2.5-ppm and 0.6-Da deviation for precursor and fragment masses, respectively; no enzyme specification; peptide ion score cutoff, 25; protein ion score cutoff, 60. Only fully tryptic peptide identifications were retained. This resulted in a high quality data set with a false positive rate of <0.2% as determined by searching randomized IPI databases with all protein sequences inverted (49). Redundancy among the identified proteins was minimized by eliminating duplicate proteins with the same Entrez gene symbol. Relative protein abundance was calculated from the ratio of numbers of peptides identified per protein. Although this provides only a rough estimate of protein abundance, it allows initial filtering of differentially expressed proteins by including those that differ >3-fold in peptide number. Protein orthologs were retrieved from Mouse Genome Informatics (MGI, www.informatics.jax.org).

Western Blotting—Protein extracts of cells were prepared as described above. Equal amounts of 25 μg of protein per sample were subjected to SDS-PAGE and immunoblotted (47) using the primary antibodies listed in Table I and HRP-coupled secondary antibodies followed by ECL.

Immunofluorescence Microscopy—Fixed cells were labeled with the primary antibodies listed in Table I and FITC- and Cy3-labeled secondary antibodies as described previously (50). Images of cells were made using a Leica TCS SP2 AOBS microscope (47) and processed with Paint Shop Pro 9.

Fluorescence-activated Cell Sorting—Cells were trypsinized, fixed, and labeled as described above using FITC or phycoerythrin-labeled secondary antibodies. At least 5,000 cells per sample were gated by FACS using a FACScan flow cytometer (BD Biosciences); the resulting data were analyzed and processed with WinMDI 2.8.

RESULTS

Identification of Proteins Specific for ESCs or Enriched in ESCs Compared with Dif-ESCs—To identify generic ESC-specific proteins and those enriched in ESCs compared with Dif-ESCs, we used human HES-2 (4) and mouse D3 (51) ESC lines. Proteins were extracted from these hESCs and mESCs as well as their differentiated derivatives obtained by culturing the cells for 12 days in the absence of feeder cells and LIF. The protein extracts of these four cell types were separated by SDS-PAGE after which the gel lanes were each sectioned into 26 gel slices. Tryptic in-gel digests were separated in 90-min capillary LC runs coupled on line to FT-ICR-MS/MS for peptide identification. Interrogation of human and mouse IPI databases resulted in the identification of ~15,000 peptides in each of the four samples. Subsequently Mascot peptide and protein cutoff scores of 25 and 60, respectively, were applied to obtain confident data sets of proteins. Redundant proteins were eliminated from the resulting data sets by removing duplicate protein identifications based on their Entrez gene number associated with each IPI entry. Hence non-redundant data sets were obtained consisting of 1,775 proteins identified in hESCs, 1,532 in differentiated hESCs (Dif-hESCs), 1,871 in mESCs, and 1,552 in differentiated mESCs (Dif-mESCs) (Supplemental Table 1, A–D, respectively). The false positive rate of these identifications was <0.2% as determined from a parallel analysis using IPI databases with all protein sequences reversed. Thus, our FT-ICR-MS/MS analysis combined with stringent cutoff filters in Mascot have resulted in high confident protein identifications.

Because our main interest was to identify proteins specifically expressed by ESCs, those proteins identified in ESCs as well as Dif-ESCs (i.e., 1,136 in human and 1,128 in mouse) (Supplemental Table 1, E and F, respectively) were eliminated from the data sets (Fig. 2A). Among the resulting 639 human and 743 mouse proteins uniquely identified in ESCs were established ESC benchmarks. These included transcription factors like Oct4 (POU5F1) (14) and UTF1 (52), telomerase-associated proteins such as Rif1 and telomere end-binding protein, and other known ESC markers (e.g., alkaline phosphatase, ALPL). The sequestration of established ESC-specific proteins in the resulting data sets substantiates our methodology to segregate ESC-specific proteins by this subtractive approach.

In addition, proteins were selected when they were more abundant in ESCs compared with Dif-ESCs based on the number of peptides identified per protein. Proteins identified by ~3 times as many peptides in ESCs than Dif-ESCs (i.e., 96 human and 145 mouse proteins) (Supplemental Table 1, G and H) were added to the subsets of proteins uniquely identified in ESCs (Supplemental Table 1, G and H). For instance, human nuclear mitotic apparatus protein 1 (NUMA1) was identified by 19 and two peptides in hESCs and Dif-hESCs, respectively, suggesting a much higher expression level in hESCs than Dif-hESCs. Likewise subunits 3, 5, and 6 of the human minichromosome maintenance (MCM) hexameric complex were identified by 6 times more peptides in hESCs than in Dif-hESCs. The peptide ratios of MCM2 and -7 were found to be slightly lower (2.8 and 3.8, respectively), whereas MCM4 was uniquely found in hESCs (identified by 18 peptides). Combined these findings suggest that the MCM complex is highly enriched in hESCs and is implicated in maintaining hESC-specific characteristics.

In total, 730 human and 888 mouse proteins were exclu-
sively found in ESCs or highly enriched in ESCs based on peptide counts (Supplemental Table 1, G and H). To further classify which of these might be specific for both human and mouse ESCs, a cross-comparison of the two data sets was made. Of the 639 putatively hESC-specific proteins, 191 matching orthologs were found among the 743 proteins uniquely identified in mESCs (Supplemental Table 2). Furthermore, 15 human and 33 mouse proteins with a peptide ratio >3-fold higher in ESCs than in Dif-ESCs were matched with mouse orthologs present in the mESC and Dif-mESC data sets. Similarly, proteins uniquely identified in mESCs (large blue circle) or >3-fold enriched in mESCs (small blue circle) were matched with human orthologs of the hESC and Dif-hESC data sets. The proteins indicated (also listed in Table I) are those for which antibodies were collected. The number of proteins in each subset are shown in red (A–C).

Fig. 2. Schematic representations of FT-ICR-MS/MS data sets of human and mouse ESCs and Dif-ESCs. A, Venn diagrams showing unique and common proteins in human (yellow) and mouse (blue) ESCs (bright) and Dif-ESCs (dark). B, Human proteins that were uniquely identified in hESCs (unique) and those with a >3-fold higher peptide ratio in hESCs compared with Dif-hESCs (>3×) were matched with mouse orthologs present in the mESC and Dif-mESC data sets (large yellow circle and small yellow circle, respectively). C, Similarly, proteins uniquely identified in mESCs (large blue circle) or >3-fold enriched in mESCs (small blue circle) were matched with human orthologs of the hESC and Dif-hESC data sets. The proteins indicated (also listed in Table I) are those for which antibodies were collected. The number of proteins in each subset are shown in red (A–C).

Western Blot Analysis of Proteins Uniquely Identified or Enriched in ESCs Compared with Dif-ESCs—Assigning proteins to the subsets described above was based on exclusive identification in either the ESC or Dif-ESC pool or the ratios of the number of peptides by which these proteins were identified in both pools. To validate this semiquantitative approach, we applied three established methods complementing the FT-ICR-MS/MS results: Western blotting, immunofluorescence confocal microscopy, and FACS cytometry (Fig. 1). We used antibodies against a selection of candidate proteins found to be either restricted to undifferentiated ESCs in our system or enriched compared with Dif-ESCs (Fig. 2C; Table I). As expected, Western blotting revealed less cytosolic β-actin (ACTB) in ESCs than in Dif-ESCs (Fig. 3A), reflecting differences in their size. Also as expected, Oct4 expression is practically restricted to ESCs (Fig. 3A), indicating that differentiation had taken place in the Dif-ESC population. Despite some variation in the protein expression ratios found with Western blotting and FT-ICR-MS/MS, the results obtained with these two methods generally coincide and confirm enrichment of candidate proteins in ESCs compared with Dif-ESCs. For example, proteasome-associated protein ECM29 homolog KIAA0368, MCM4, and DNA topoisomerase IIα (TOP2A) were all uniquely identified in hESCs by FT-ICR-MS/MS. The Western blots accordingly show that these proteins are specifically expressed by hESCs (e.g. TOP2A) or are highly enriched in hESCs (e.g. KIAA0368 and MCM4) (Fig. 3A). Similarly expression levels of proteins found exclusively in mESCs, like MSH6, or identified with a >10-fold peptide ratio, like calcyclin-binding protein (CACYBP) and X-linked ribosomal protein S4 (RPS4X), are in line with the Western blots (Fig. 3A).

Immunofluorescence Microscopic Imaging of Proteins Uniquely Identified or Enriched in ESCs Compared with Dif-ESCs—To determine ESC specificity of candidate proteins at
In the cellular level, we double stained cells with anti-Oct4/FITC-labeled secondary antibodies and antibodies recognizing the candidate ESC proteins (listed in Table I) with Cy3-labeled secondary antibodies. Confocal laser scanning microscopy showed that the vast majority of cells in human and mouse ESC colonies were positive for Oct4 (Supplemental Fig. 1A).

### TABLE II

| Gene symbol | IPI accession no. | Protein name, human | No. of peptides | Mascot score | Peptide ratio, mouse |
|-------------|------------------|---------------------|----------------|-------------|---------------------|
| Human       | Mouse            |                     | Human          | Mouse       |                     |
| ACAC1       | IPI00121833      | 3-Ketoacyl-CoA thiolase | 10             | 61          | 415                | 3.3               |
| ACTL6A      | IPI000323660     | Actin-like protein 6A | 5              | 5           | 289                | 5.0               |
| AD-003      | IPI00153664      | Adrenal gland protein AD-003 | 4              | 4           | 221                | 4.0               |
| AHSA1       | IPI00153740      | Activator of 90-kDa heat shock protein ATPase homolog 1 | 5              | 8           | 191                | 344               |
| ARS2        | IPI00224643      | Arsenite resistance protein 2 | 5              | 6           | 220                | 278               |
| BCAS2       | IPI000316184     | Putative spliceosome-associated protein | 2              | 6           | 65                 | 283               |
| BZW1        | IPI00132588      | Basic leucine zipper and W2 domains 1 | 10             | 10          | 432                | 397               |
| C10orf119   | IPI000153724     | Hypothetical protein | 2              | 3           | 111                | 166               |
| CNAP1       | IPI00172226      | Condensin subunit 1 | 7              | 6           | 286                | 188               |
| CSIG        | IPI0008708       | Cellular senescence-inhibited gene protein | 3              | 9           | 122                | 474               |
| DNAJC7      | IPI00031385      | DNAJ (Hsp40) homolog subfamily C member 7 | 4              | 7           | 200                | 392               |
| EIF2S2      | IPI00116302      | Eukaryotic translation initiation factor 2 subunit 2 | 2              | 7           | 96                 | 292               |
| GA17        | IPI00115580      | DNA mismatch repair protein MSH6 | 15             | 25          | 585                | 1002              |
| HAT1        | IPI00115597      | Methyl-CpG-binding protein 3 | 5              | 4           | 250                | 258               |
| MBD3        | IPI00131067      | NAD-dependent malic enzyme | 3              | 5           | 187                | 271               |
| ME2         | IPI00011201      | Histone acetyltransferase type B catalytic subunit | 4              | 9           | 166                | 426               |
| MSH6        | IPI00132576      | Small nuclear ribonucleoprotein polypeptide B | 4              | 6           | 171                | 282               |
| MYBBP1A     | IPI0005024       | MYB-binding protein 1A | 11             | 24          | 511                | 1257              |
| NDUF81      | IPI00121833      | NADH dehydrogenase (ubiquinone) 1β subcomplex 10 | 5              | 8           | 277                | 341               |
| NDUF81      | IPI00187677      | NADH dehydrogenase (ubiquinone) Fe-S protein 1 | 5              | 8           | 277                | 341               |
| NS          | IPI0006451       | Nucleostemin | 2              | 8           | 130                | 336               |
| NSF         | IPI00031934      | Vesicle-fusing ATPase | 2              | 6           | 107                | 301               |
| NUP153      | IPI00132576      | NUP153 | 3              | 5           | 147                | 183               |
| PES1        | IPI00011042      | Pescadillo homolog 1 | 3              | 6           | 159                | 308               |
| PPAT        | IPI00228949      | Amidophosphoribosyltransferase precursor | 3              | 6           | 142                | 316               |
| ppm1g       | IPI000306882     | Protein phosphatase 2C γ isoform | 5              | 8           | 328                | 425               |
| PPP2R2A     | IPI00462007      | Serine/threonine protein phosphatase 2A | 2              | 7           | 67                 | 372               |
| RFC5        | IPI00132481      | Replication factor C (activator 1) | 1              | 9           | 64                 | 418               |
| SDHA        | IPI00132576      | Solute carrier family 25 member 3 | 6              | 13          | 277                | 615               |
| SLC25A3     | IPI00230351      | Succinate dehydrogenase (ubiquinone) flavoprotein subunit | 2              | 12          | 135                | 558               |
| SNRPB2      | IPI00132576      | Small nuclear ribonucleoprotein polypeptide B | 4              | 6           | 171                | 282               |
| TOP2A       | IPI00122223      | DNA topoisomerase IIα | 20             | 30          | 864                | 1339              |
| ZNF9        | IPI00430812      | Cellular nucleic acid-binding protein | 6              | 10          | 241                | 404               |
Undifferentiated hESCs were predominantly localized in the outer ring extending to the edge of each colony surrounding a relatively small cluster of Oct4-negative cells localized in the center. On the other hand, undifferentiated Oct4-positive mESCs grew as dense colonies of various size with only a few differentiated Oct4-negative mouse cells scattered between the colonies.

Expression of the candidate proteins was mainly restricted to ESCs positive for Oct4 (Supplemental Fig. 1, B–Q). For example, MCM4 was clearly present in mESCs (Fig. 4B) but not in Dif-mESCs (Fig. 4A). Although less pronounced, a similar difference in MCM4 expression was observed between mESC and Dif-mESCs (Supplemental Fig. 1H). This results support the FT-ICR-MS/MS data (Table I) and coincide with the Western blots (Fig. 3). On the other hand, CACYBP was barely detectable in hESCs or Dif-hESCs (Supplemental Fig. 1A). This once again confirmed the conclusion based on the Western blots (Fig. 3A) and the ESC to Dif-ESC peptide ratio of mouse cells (i.e. 4.50) that is ~3.5-fold higher than that of human cells (i.e. 1.29) (Table I). Comparative analysis showed an even more pronounced difference for RPS4X with a peptide ratio of 1.7 for human cells and 19.5 for mouse cells (Fig. 3A and Supplemental Fig. 1N). Thus, although some of the proteins are not restricted to ESCs, their expression levels are higher in ESCs than in Dif-ESCs and appear to be species-dependent.

Interestingly MCM4 (Fig. 4A), karyopherin-α2 (KPNA2) (Fig. 4E), and TOP2A (Fig. 4F) in HES-2 ESCs as well as CACYBP (Fig. 4B) and MSH6 (Fig. 4G) in mESCs showed highly variable expression in individual cells within the same colony. Stages of the cell cycle and differentiation will affect processes like genome replication, which may be relevant for MCM4, MSH6, and TOP2A. Of note were small cell clusters among the Dif-ESC populations of human and mouse cells that expressed methylenetetrahydrofolate dehydrogenase 1 (MTHFD1) (Fig. 4H) and MSH6 (Fig. 4f), respectively. Co-localization with Oct4 suggested that these cells retained ESC characteristics despite 12 days of culture in the absence of ESC-sustaining factors. Taken together, these findings indicate that a relatively high number of ESCs expressed the candidate proteins at substantially higher levels than Dif-ESCs, corroborating the FT-ICR-MS/MS and Western blot data at the cellular level.

FACS-mediated Segregation of Human and Mouse ESCs from Dif-ESCs—Because immunofluorescence analysis by confocal microscopy is limited to subpopulations of individual cells, we extended the study with FACS cytometry to test whether the antibodies with high specificity could be used to
segregate ESCs from Dif-ESCs. Approximately 82 and 95% of cells in the hESC and mESC cultures, respectively, reached arbitrary thresholds of Oct4-dependent immunofluorescence intensity (Fig. 5, A and B, respectively) that were calibrated with that of their respective Dif-ESCs. Notably ~7% of the Dif-hESCs expressed Oct4 to the same extent after 12 days of differentiation, and this was almost 2-fold higher for the pool of Dif-mESCs (~13%). The Oct4-positive cells among Dif-mESCs are also evident as a faint band on the Western blot (Fig. 3A). Although less than Oct4, both KPNA2 and MSH6 appeared relatively specific for hESCs; these proteins gated ~77% (Fig. 5C) and ~49% (Fig. 5D) of the hESC population, respectively, approximately twice as many as Dif-hESCs (~38 and ~22%, respectively). On the other hand, MSH6 is highly specific for mESCs (Fig. 3E), sorting ~92% of mESCs and ~13% of Dif-mESCs. These numbers closely match that of Oct4 (above) and UTF1, sorting ~92% of the mESC and ~11% of the Dif-mESC population (Fig. 5F). Thus, under these conditions, MSH6 can be considered as a specific marker for mESCs.

Expression Comparison of Putative hESC-specific Proteins in Three Different hESC Lines: HES-2, HUES-1, and NL-HESC-01—To investigate whether the selected set of proteins markedly enriched in HES-2 hESCs, as validated by Western blotting and immunofluorescence microscopy, could serve as generic markers for hESCs, two other hESC lines, HUES-1 (5) and NL-HESC-01 (53), were also analyzed by Western blotting. In contrast to HES-2 hESCs, which are passaged mechanically by a cut-and-paste method in serum-containing medium (4, 45), HUES-1 hESCs are passaged enzymatically by trypsinization and cultured in serum replacement with basic fibroblast growth factor (5). HUES-1 hESCs were differentiated either in recommended HUES-1 differentiation medium (5) or that used for HES-2 cells (4, 45). Expression of β-actin and Oct4 in HUES-1 hESCs and Dif-hESCs resembled that of corresponding HES-2 cells, but the expression ratios differed for several other proteins in HUES-1 cells (Fig. 3B). For example, the expression of KIAA0368 as well as KPNA2 and Ran-binding protein 2 (RANBP2) was down-regulated upon differentiation of HES-2 cells but was unchanged in HUES-1 cells (Fig. 3B), independent of differentiation conditions. This could be due to differences in the rate or type of differentiation or in the original hESC isolation methods (for a review, see Ref. 54). By contrast, expression of MSH6 was reduced to a greater extent in HUES-1 cells differentiated in HUES-1 differentiation medium (5) or that used for HES-2 cells (4, 45). Expression of β-actin and Oct4 in HUES-1 hESCs and Dif-hESCs resembled that of corresponding HES-2 cells, but the expression ratios differed for several other proteins in HUES-1 cells (Fig. 3B).

We also analyzed our new hESC line, NL-HESC-01 (53), derived and passaged as HES-2 hESCs but differentiated by aggregation in suspension culture for 5 days followed by 7-day attached growth. Cell lysates from NL-HESC-01 hESCs of passage 22 and passage 56 as well as NL-HESC-01 Dif-hESCs of passage 56 were analyzed for differential expression of candidate proteins by Western blotting. In contrast to HES-2 and HUES-1, the expression of β-actin is higher in NL-HESC-01 hESCs than in Dif-hESCs (Fig. 3B), possibly
reflecting the difference in differentiation method. With the exception of RANBP2 and MSH6, the overall expression of candidate proteins by NL-HESC-01 hESCs resembled HES-2 more closely than HUES-1 hESCs, irrespective of passage number. These findings suggest that hESCs cultured under identical conditions (i.e., cut-and-paste and serum-containing medium) have similar protein expression profiles, independent of their genetic origin. Notably NL-HESC-01 hESCs maintain this expression profile through multiple passages. Combined with previous findings (55, 56), these results suggest

![Confocal laser scanning microscopic images of human and mouse ESCs and Dif-ESCs.](image)

**Fig. 4.** Confocal laser scanning microscopic images of human and mouse ESCs and Dif-ESCs. Fixed cells were incubated with primary antibodies as indicated and secondary antibodies fluorescently labeled with FITC (green) or Cy3 (red). Simultaneous visualization of Oct4 and MCM4 in hESCs (A) and Dif-hESCs (B), Oct4 and CACYBP in mESCs (C) and in Dif-mESCs (D), KPNA2 (E) and TOP2A (F) in hESCs, MSH6 in mESCs (G), Oct4 and MTHFD1 in Dif-hESCs (H), and Oct4 and MSH6 in Dif-mESCs (I) is shown. Oct4-positive cells in the Dif-ESC populations are indicated with arrowheads (H and I). ESCs formed separate colonies, and Dif-ESCs developed as confluent layers. Scale bars, 20 μm.
that mechanical passaging not only supports karyotypic stability but also sustains a stable protein expression profile. Interestingly TOP2A is detected in all three hESC lines but not in their respective differentiated derivatives (Fig. 3, A and B), indicating that this protein is a putative marker of hESCs like MSH6 and possibly MTHFD1 for mESCs (Fig. 3A).

DISCUSSION

The power of FT-ICR-MS/MS is its sensitivity and exceptional mass accuracy in the detection and identification of proteins. Coupled to protein and peptide separation techniques, we used large scale comparative FT-ICR-MS/MS analysis as a rapid high throughput screen for proteins associated with ESCs. This strategy resulted in the identification of 191 proteins associated with ESCs of both species (Supplemental Table 2). The reduction of ~1,800 proteins initially identified in ESCs to 191 potentially generic ESC-specific proteins can now be used for further biological validation in a more targeted approach. Our approach proved highly effective as a relatively fast search for ESC-specific proteins on a large scale. The sensitivity of our methodology is illustrated by the identification of several low abundance transcription factors like Oct4 and UTF1, which are specific for ESCs.

To determine altered protein expression levels in ESCs after differentiation, we conducted a semiquantitative analysis based on peptide count, which was validated for a subset of proteins using conventional analytical techniques with antibodies. The peptide ratios determined for the selected proteins generally corresponded to the protein levels observed on Western blot and additionally unveiled discrepancies between human and mouse ESCs. For instance, mouse RANBP2 is not visible on Western blot (Fig. 3A), whereas the immunofluorescence images show obvious differences between mESCs and Dif-mESCs (Supplemental Fig. 1N). This suggests that the antibody does not recognize the denatured form of its target protein. The ratio between MCM4 expression in mESCs and Dif-mESCs appears less pronounced on Western blot (Fig. 3A) compared with immunofluorescence microscopy (Supplemental Fig. 1I). However, Western blot analysis indicates that the antibody binds to several additional antigens with molecular masses lower than the ~90 kDa of MCM4 (data not shown) that could be degradation products of MCM4 that bias actual protein ratios. In addition, the immunofluorescence images of hESCs incubated with antibodies against MTHFD1, RPS4X, and USP9X (Supplemental Fig. 1, K, O, and Q, respectively) showed that these proteins may also be expressed by the surrounding feeder cells. Precise excision of the hESCs from the co-cultures for FT-ICR-MS/MS and Western blot analysis ensured that contamination with non-hESC proteins was minimal.

A similar excision method prior to FT-ICR-MS/MS and Western blot analysis could not be applied to the Dif-ESC pools for exclusion of cells that maintained characteristics of ESCs (e.g. Oct4 expression); the cells cannot be distinguished without specific labeling. The FACS data indicate that the proportion of ESCs among the Dif-ESC population is higher in mouse than human cultures (Fig. 5, compare A and B). Therefore, proteins found to be unique in hESCs yet designated as...
highly enriched in mESCs (i.e., present in the Dif-mESC pool albeit at very low levels, e.g., MSH6 and TOP2A) (Supplemental Table 1G) may actually be very specific for mESCs. This notion is supported by the confocal image of MSH6-positive cells among Dif-mESCs as these cells also express Oct4 (Fig. 4B).

As proof of principle for our methodological approach, a number of antibodies was used to validate a selection of proteins differentially identified in ESCs and Dif-ESCs by FT-ICR-MS/MS. Although we could not confirm the expression of all proteins identified because of limited availability of antibodies, the data sets represent the highest number of putatively ESC-specific or enriched proteins described to date using any proteomic method and the first on hESCs. Of particular interest are the uncharacterized proteins among the 191 proteins found exclusively in ESCs (Supplemental Table 2) as these may include novel ESC-specific benchmarks and provide new clues to explain species-related differences and similarities of ESCs. The reliability with which proteins were assigned to the specified subsets according to their peptide ratios was determined by calculating the percentage of proteins that would be allocated to the same subsets based on the Western blot results. We obtained 40 antibodies against 26 different proteins that FT-ICR-MS/MS data indicated were enriched in ESCs compared with Dif-ESCs. Expression was analyzed by Western blotting using cell lysates of ESCs and Dif-ESCs derived from both species. In total, 18 of the 26 human and 16 of the 26 mouse proteins were detected by their respective antibodies. Of the 18 human and 16 mouse proteins detected in the cell lysates, 15 human (85%) and 12 mouse (75%) proteins, respectively, showed significantly increased expression in ESCs compared with Dif-ESCs (Fig. 3A), supporting the FT-ICR-MS/MS findings. Generating antibodies against unknown proteins among the 191 identified in human and mouse ESCs may provide a set of additional tools for confirming identity, monitoring hESC and mESC characteristics in culture, and possibly predicting the differentiation potential of newly derived lines.

It is not surprising that many of the validated proteins found to be restricted to or highly enriched in ESCs are involved in cell cycle progression as these cells proliferate at relatively high rates. For example, proteins for DNA replication, such as human TOP2A and mouse MSH6, are specifically expressed by ESCs on Western blot in the present study (Fig. 3A) but may also be active in other cells that proliferate. This suggests that cell cycle-related proteins per se cannot be used as ESC-specific markers. However, FACS with the antibody against MSH6 gated UTF1-positive mESCs (data not shown) with an efficacy similar to Oct4-based sorting (Fig. 5B). Moreover, the antibody effectively labeled undifferentiated (i.e., UTF1-positive) cells that were present among the differentiated mESC population despite the absence of ESC-sustaining conditions. Thus, differences in the overall expression level of proteins as such between ESCs and Dif-ESCs, whether they are cell cycle-related or not, may be sufficient to label and isolate ESCs from mixed populations, indicating they can be used as ESC-specific selectors under these conditions. In addition, they collectively provide insight into variations in cell cycle progression or proliferative capacity of different ESC lines under certain growth conditions providing ESC line-specific fingerprints.

Although species-specific protein profiles of ESCs were expected to vary to some extent, differences between ESC lines derived from the same species were also observed. This is exemplified by a comparison of our mESC protein data set with those previously published (24, 25). Although 194 of the 218 proteins (89%) identified in R1 mESCs (24) were also present in our data set of D3 mESCs, the former study described a relatively small number of proteins that is probably dominated by those abundantly expressed. On the other hand, of the 1,790 proteins identified in E14-1 mESCs (25), we detected 842 (47%) in D3 mESCs. Moreover 65 of these proteins were among the subset of 191 proteins (34%) uniquely identified in both human and mouse ESCs and thus are likely to be involved in ESC maintenance.

Comparative microarray studies have repeatedly shown divergent gene expression profiles of multiple ESC lines established and cultured in different laboratories even when derived from the same species (21–23). Discrepancies in expression patterns have been explained by (i) differences in culture and derivation procedures applied in laboratories, (ii) variations in analysis techniques, and (iii) diverse genetic backgrounds of the multiple ESC lines established to date. Even with relatively small data sets distinct differential protein expression profiles of HUES-1 and NL-HESC-01 cells became apparent on Western blot (Fig. 3B). In contrast to HUES-1 hESCs, which are passaged enzymatically, the overall protein pattern of NL-HESC-01 hESCs appeared fairly similar to that of HES-2 hESCs (Fig. 3A); both of these cell lines were passaged mechanically. These observations suggest that variations in proteomic profiles of established hESC lines may be induced by differences in culture techniques and preceding derivation procedures. In addition, aberrant protein expression could result from genetic disparities; it has been shown recently that genomic alterations may occur and tend to accrue over extended culturing periods (56).

Collectively our findings indicate that mass spectrometric techniques have now matured sufficiently to be applied to wide scale identification of putative benchmarks, which complement microarray studies and may provide the missing link between gene transcription and cell behavior. Further analysis of the listed candidates (Supplemental Table 2) at the functional level may very well uncover novel generic ESC-specific benchmarks that are essential for maintaining the undifferentiated state of ESCs. In addition, the small scale comparison of three individual hESC lines displays indisputable differences that most likely result from variations in culture methods. These preliminary findings emphasize the importance of
extending genomic and transcriptomic characterization of existing and novel hESC lines (26, 27) with proteomic profiling, which is best addressed using mass spectrometry to achieve the necessary scale. This assimilation of basic data to identify the factors that determine the gene and protein expression profiles of hESCs is an essential prelude to their clinical application in the light of their potential instability and ability to maintain pluripotency in culture.

Acknowledgments—We thank D. A. Melton for HUES1 hESCs, M. F. Peres for HES2-2 hESCs, S. Van den Brink for help with the NL-HESC-01 hESC and D3 mESC cultures, J. Kuipers for technical assistance with the FACS, K. W. Rodenburg for cold water fish gelatin, and the authors listed in Table I for sending the antibodies.

* This work was supported by the Netherlands Proteomics Centre Bisk program (to D. V. H., A. J. R. H., C. L. M., and J. K.), the inter-University Cardiology Institute of the Netherlands (to R. P., D. W.-V. O., and C. L. M.), and ES Cell International Pte. Ltd. (to C. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

[5] The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

To whom correspondence may be addressed. Tel.: 31-30-212-1800; Fax: 31-30-251-8219; E-mail: christin@niob.knaw.nl.

** To whom correspondence may be addressed. Tel.: 31-30-253-9974; Fax: 31-30-251-6464; E-mail: j.krijgsveeld@chem.uu.nl.

REFERENCES

1. Evans, M. J., and Kaufman, M. H. (1981) Establishment in culture of pluripotent cells from mouse embryos. Nature 292, 154–156
2. Martin, G. R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. 78, 7634–7638
3. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282, 1145–1147
4. Neuhaus, B. E., Peras, M. F., Fong, C. Y., Trounson, A., and Bongso, A. (2000) Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. 18, 399–404
5. Cowan, C. A., Klimanskaya, I., McMahon, J., Atienza, J., Wmitry, J., Zucker, J. P., Wang, S., Morton, C. C., McMahon, A. P., Powers, D., and Melton, D. A. (2004) Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. 350, 1353–1356
6. Keller, G. (2000) Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev. 19, 1129–1155
7. Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature 309, 255–256
8. Marshak, D. R., Gardner, R. L., and Gottlieb, D. (2001) Embryonic Stem Cells in Stem Cell Biology, pp. 205–230, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
9. Peras, M. F., and Trounson, A. O. (2004) Human embryonic stem cells: prospects for development. Development 131, 5515–5525
10. Wobus, A. M., and Boheler, K. R. (2005) Embryonic stem cells: prospects for developmental biology and cell therapy. Physiol. Rev. 85, 635–678
11. Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Sough, N. M. (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336, 684–687
12. Ying, Q. L., Nichols, J., Chambers, I., and Smith, A. (2003) BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115, 281–292
13. Boiani, M., and Scholer, H. R. (2005) Regulatory networks in embryono-
34. Imamoto, N., Shimamoto, T., Takao, T., Tachibana, T., Kose, S., Matsubae, M., Sekimoto, T., Shimonishi, Y., and Yoneda, Y. (1995) In vivo evidence for involvement of a 58 kDa component of nuclear pore-targeting complex in nuclear protein import. EMBO J. 14, 3617–3626

35. Ishimi, Y., Komamura-Kohno, Y., Karasawa-Shimizu, K., and Yamada, K. Levels of MCM4 phosphorylation and DNA synthesis in DNA replication block checkpoint control. J. Struct. Biol. 146, 234–241

36. Hernandez-Pigeon, H., Laurent, G., Humbert, O., Salles, B., and Lautier, D. (2004) Degradation of mismatch repair in MMR-deficient cells involving the ubiquitin-proteasome pathway. FEBS Lett. 562, 40–44

37. Brody, L. C., Conley, M., Cox, C., Kirke, P. N., McKeever, M. P., Mills, J. L., Molloy, A. M., O’Leary, V. B., Parle-McDermott, A., Scott, J. M., and Swanson, D. A. (2002) A polymorphism in R65Q, in the trinucleosomal enzyme methylenetetrahydrofolate dehydrogenase/methylenetetrahydrofolate cyclohydrolase/formyltetrahydrofolate synthetase is a maternal genetic risk factor for neural tube defects: report of the Birth Defects Research Group. Am. J. Hum. Genet. 71, 1207–1215

38. Ferro, E. S., Hyslop, S., and Camargo, A. C. (2004) Intracellular peptides as putative natural regulators of protein interactions. J. Neurochem. 91, 769–777

39. Guthridge, M. A., Bellosta, P., Tavoloni, N., and Basilio, C. (1997) FIN13, a novel growth factor-inducible serine-threonine phosphatase which can inhibit cell cycle progression. Mol. Cell. Biol. 17, 5485–5498

40. Pichler, A., Gast, A., Seeler, J. S., Dejean, A., and Melchior, F. (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell 108, 109–120

41. Zinn, A. R., Alagappan, R. K., Brown, L. G., Wool, I., and Page, D. C. (1994) The FAM deubiquitylating enzyme localizes to multiple points of protein trafficking in epithelia, with clinical outcome. J. Struct. Biol. 115, 4001–4012

42. Mummery, C. L., Slager, H., Krijger, W., Feijen, A., Freund, E., Koornneef, I., and Van den Eijnden-van Raaij, A. J. (1990) Expression of transforming growth factor β2 during the differentiation of murine embryonal carcinoma and embryonic stem cells. Dev. Biol. 137, 161–170

43. Van Hoof, D., Rodenburg, K. W., and Van der Horst, D. J. (2002) Insect lipoprotein follows a transferrin-like recycling pathway that is mediated by the insect LDL receptor homologue. J. Cell Sci. 115, 4001–4012

44. Murning, H. D., Van der Heeft, E., Ten Hove, G. J., and De Jong, A. P. J. M. (2002) Nanoscale LC-MS(n): technical design and applications to peptide and protein analysis. J. Sep. Sci. 25, 557–568

45. Peng, J., Elias, J. E., Thoren, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. 2, 43–50

46. Van Hoof, D., Rodenburg, K. W., and Van der Horst, D. J. (2005) Receptor-mediated endocytosis and intracellular trafficking of lipoproteins and transferrin in insect cells. Insect Biochem. Mol. Biol. 35, 117–128

47. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985) The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J. Embryol. Exp. Morphol. 87, 27–45

48. Okuda, A., Fukushima, A., Nishimoto, M., Orimo, A., Yamagishi, T., Nabeshima, Y., Kuro-o, M., Nabeshima, Y., Boon, K., Keaveney, M., Stunnenberg, H. G., and Muramatsu, M. (1998) UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra-embryonic cells. EMBO J. 17, 2019–2032

49. Van de Stolpe, A., Van den Brink, S., Van Rooijen, M., Ward-van Oostwaard, D., Van Inzen, W., Slaper-Cortenbach, I., Fauser, B., Van den Hout, N., Weima, S., Passier, R., Smith, N., Denning, C., and Mummery, C. (2005) Human embryonic stem cells: towards therapies for cardiac disease. Derivation of a Dutch human embryonic stem cell line. Reprod. Biomed. Online 11, 476–485

50. Hoffmann, L. M., and Carpenter, M. K. (2005) Characterization and culture of human embryonic stem cells. Nat. Biotechnol. 23, 699–708

51. Buzzard, J. J., Gough, N. M., Crook, J. M., and Colman, A. (2004) Karyotype analysis of human embryonic stem cells. Nat. Biotechnol. 22, 381–382

52. Maitra, A., Arking, D. E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D. J., Liu, Y., Brimble, S. N., Noaksson, K., Hyllner, J., Schulz, T. C., Zeng, X., Freed, W. J., Crook, J., Abraham, S., Colman, A., Sartipy, P., Matsui, S., Carpenter, M., Gazdar, A. F., Rao, M., and Chakravarti, P. (2005) Genomic alterations in cultured human embryonic stem cells. Nat. Genet. 37, 1099–1103