Quantitative Proteomics Analysis Demonstrates Post-transcriptional Regulation of Embryonic Stem Cell Differentiation to Hematopoiesis*⑤

Andrew J. K. Williamson‡, Duncan L. Smith§, David Blinco‡, Richard D. Unwin‡, Stella Pearson§, Claire Wilson§, Crispin Miller§, Lee Lancashire§, Georges Lacaud§, Valerie Kouskoff§, and Anthony D. Whetton‡¶

Embryonic stem (ES) cells can differentiate *in vitro* to produce the endothelial and hematopoietic precursor, the hemangioblasts, which are derived from the mesoderm germ layer. Differentiation of $Bry^{GFP/+}$ ES cell to hemangioblasts can be followed by the expression of the $Bry^{GFP/+}$ and *Flk1* genes. Proteomic and transcriptomic changes during this differentiation process were analyzed to identify mechanisms for phenotypic change during early differentiation. Three populations of differentiating Bry^{GFP} ES cells were obtained by flow cytometric sorting, GFP⁻Flk1⁻ (epiblast), GFP⁺Flk1⁻ (mesoderm), and GFP⁺Flk1⁺ (hemangioblast). Microarray analyses and relative quantification two-dimensional LCLC-MS/MS on nuclear extracts were performed. We identified and quantified 2389 proteins, 1057 of which were associated to their microarray probe set. These included a variety of low abundance transcription factors, *e.g.* UTF1, Sox2, Oct4, and E2F4, demonstrating a high level of proteomic penetrance. When paired comparisons of changes in the mRNA and protein expression levels were performed low levels of correlation were found. A strong correlation between isobaric tag-derived relative quantification and Western blot analysis was found for a number of nuclear proteins. Pathway and ontology analysis identified proteins known to be involved in the regulation of stem cell differentiation, and proteins with no described function in early ES cell development were also shown to change markedly at the proteome level only. ES cell development is regulated at the mRNA and protein level. *Molecular & Cellular Proteomics 7:459–472, 2008.*

In the mouse embryo the process of gastrulation generates the mesodermal layer. During embryo development the mesoderm is generated from the epiblast or embryonic ectoderm.

The cells involved in mesodermal commitment show a marked increase in Brachyury (*Bry*)[1] expression (1). The mesoderm migrates and develops into tissues including the endothelium as well as blood (2, 3). A common progenitor for these cell types, the hemangioblast, is first detected at the midstreak stage of gastrulation (4). These processes of early cell development within the embryo are not easily assessed because of limitations on the material available for study. This can be overcome by using embryonic stem (ES) cell lines derived from the inner cell mass of the blastocyst-stage embryo (5). ES cells can be allowed to differentiate *in vitro* to form three-dimensional spheroid cultures known as embryoid bodies (EBs). These structures contain derivatives of all three germ cell layers: mesoderm, endoderm, and ectoderm (6, 7). Mesoderm-derived populations contained within the EBs include hematopoietic and endothelial progeny. This model has been likened to the temporal pattern of development of these populations in embryogenesis (8–10), and EB cultures have been shown to contain cells with hemangioblast potential (11, 12).

Differentiating ES cells *in vitro* with the use of the EB model system have been valuable in identifying a number of genes associated with self-renewal, such as *Oct4*, *Sox2*, and *Nanog* (13–16). These proteins are considered as markers for the pluripotent embryonic stem cell in mouse and human (13, 17–20). Disruption of *Oct4* or *Nanog* results in the differentiation of ES cells to trophectoderm and extraembryonic endoderm, respectively (13, 17, 18, 20). The target genes for these proteins often encode transcription factors and chromatin-remodeling enzymes (21). Other genes have key roles in the process of mesoderm and hemangioblast development.

[1] The abbreviations used are: Bry, Brachyury; ES, embryonic stem; GFP, green fluorescent protein; $Bry^{GFP/+}$ ES, GFP has been knocked into one locus of the Brachyury gene, replacing exon 1; EB, embryoid body; Flk1, vascular endothelial growth factor 1 receptor 2; HMG, high mobility group; HDAC, histone deacetylase; DNMT1, DNA methyltransferase 1; iTRAQ, isobaric tags for relative and absolute quantitation; LIF, leukemia inhibitory factor; SOM, self-organizing map; MTG, monothioglycerol; GO, gene ontology; Rb, retinoblastoma.

For example the transcription factors Scl/tal-1 and Runx1 have been shown to function in early development specifically at the stage of hematopoietic commitment (22–26).

Bry is a member of the T-box gene family encoding transcription factors that are thought to be involved in differentiation (27). The majority of the cells contained within the primitive streak are *Bry*-positive; expression is rapidly downregulated as the cells migrate (1). *In vitro* Fehling *et al.* (28) have targeted the *Bry* locus with green fluorescent protein (*GFP*) and demonstrated that it was expressed in mesoderm; this was later confirmed to be the case *in vivo* (4). *Flk1*, the vascular endothelial growth factor 1 receptor 2, is essential for yolk sac blood island development and vasculogenesis; knockout experiments demonstrated reduced hematopoiesis (29–31). The process of mesodermal precursor cell generation within the primitive streak and migration to the site of yolk sac development is associated with *Flk1* expression (31). In differentiating ES cells the expression of *Flk1* overlaps that of *Bry* and the formation of the hemangioblast-like cells (23). It was subsequently shown that identifying *Bry* and *Flk1* expression enabled the enrichment of three populations of cultured ES cells. These represent a premesoderm stage (Bry$^-$Flk1$^-$), a prehemangioblast mesoderm stage (Bry$^+$Flk1$^-$ cells), and the hemangioblast cell stage (Bry$^+$Flk1$^+$). Gene expression analysis of the three populations of cells showed striking differences that were linked to the biological potential of the cells. For example two members of the Wnt family, *Wnt3a* and *Wnt8a*, whose expression is seen in the primitive streak and early mesoderm, were only seen in GFP$^-$Flk1$^-$ and GFP$^+$Flk1$^-$ cells (28) (also see Fig. 1). Runx1 and Scl were predominantly seen in the GFP$^+$Flk1$^+$ cells, which is consistent with the ability of this population of cells to generate hemangioblast cells (28).

Protein families such as transcription factors and chromatin-remodeling enzymes are known to be the target of post-transcriptional regulation that affects the level of protein, whereas mRNA levels are unchanged (32–35). Using systematic approaches mRNA levels have been shown in several systems not to comprehensively reflect changes in protein levels (36–41). This observation argues that systematic analyses of stem cells and progenitor cells must include proteomics analysis for full details to be acquired of regulatory pathways associated with differentiation.

We have therefore established a method for defining the proteome of stem cell populations using isobaric tags for relative and absolute quantitation (iTRAQ). This allows four samples to be analyzed simultaneously, giving relative quantification on many hundreds of proteins from a relatively small (1–5 million) cell number. Unwin *et al.* (42) compared transcriptome changes with those in the proteome using hematopoietic stem cell-enriched populations compared with their differentiated progeny. They demonstrated that 54% of protein changes observed are not seen at the transcriptional level. Thus the process of differentiation is governed at the protein turnover/synthesis level as well as by rates of mRNA synthesis/degradation. To obtain the detail on nuclear protein regulation in enriched populations of primitive cells we used premesoderm stage (Bry$^-$Flk1$^-$) cells, a prehemangioblast mesoderm stage (Bry$^+$Flk1$^-$), and hemangioblast stage (Bry$^+$Flk1$^+$) cells. These are enriched from differentiating ES cells where *GFP* is targeted to the *Bry* locus. This offers an effective marker for the mesodermal population that develops within the EBs. Here we systematically study this progression in development and show that protein expression is controlled at the post-transcriptional and translational levels in developing ES cells.

EXPERIMENTAL PROCEDURES

Cell Culture—The *GFP* gene was knocked into one allele of the *Bry* gene of the mouse embryonic stem cell line E14.1 replacing exon 1 as described in Fehling *et al.* (28) generating the *Bry*$^{GFP/+}$ ES cell line. Maintenance and differentiation of these cells have been described previously (28, 43). Briefly ES cells were maintained on irradiated embryonic feeder cells in Dulbecco's modified Eagle's medium supplemented with 15% (v/v) FCS, penicillin, streptomycin, LIF (1% (v/v) conditioned medium (44, 45)), and 1.5×10^{-4} M monothioglycerol (MTG; Sigma). Prior to the onset of differentiation the cells were passaged twice on gelatinized plates. The first passage was into Dulbecco's modified Eagle's medium supplemented as above, and the second was into Iscove's modified Dulbecco's medium supplemented with 15% (v/v) FCS, penicillin, streptomycin, LIF (1% (v/v) conditioned medium), and 1.5×10^{-4} M MTG (Sigma). The ES cell population was harvested at this stage.

For the generation of EBs, ES cells were trypsinized and plated at 2.5×10^4 cells/ml. Differentiation of EBs was carried out in 100-mm Petri grade dishes in Iscove's modified Dulbecco's medium supplemented with 15% (v/v) FCS, 2 mM L-glutamine (Invitrogen), 200 μg/ml transferrin, 0.5 mM ascorbic acid (Sigma), and 4.5×10^{-4} M MTG.

Flow Cytometry and Cell Sorting—EBs were harvested and trypsinized, and the resultant single cell suspension was analyzed on a FACSCalibur flow cytometer (BD Biosciences). Staining and isolation of cell populations were performed as described previously (28). Co-expression of GFP with the receptor tyrosine kinase *Flk1* revealed the emergence of three distinct cell populations, GFP$^-$Flk1$^-$, GFP$^+$Flk1$^-$, and GFP$^+$Flk1$^+$ cells, which represent a developmental progression ranging from premesoderm to prehemangioblast mesoderm to the hemangioblast. These sorted populations were used for analysis of the proteome and microarray analysis of transcription. Furthermore the gene expression pattern of each sorted population was assessed to ensure that appropriate cell enrichment had occurred.

Gene Expression Analysis—Total RNA was extracted using an RNeasy minikit and treated with RNase-free DNase (Qiagen, Valencia, CA). Total RNA (2 μg) was reverse transcribed into cDNA with random hexamers using the Omniscript RT kit (Qiagen). PCR was performed with ~100 ng of cDNA and Biomix Red (Bioline). Cycling conditions were as follows: 95 °C for 3 min followed by 25–35 cycles of amplification (95 °C denaturation for 30 s, annealing for 30 s, and 72 °C elongation for 30 s) with a final incubation at 72 °C for 5 min. Details of primer sequences, annealing temperature, and cycle numbers for each PCR are shown in Supplemental Table 1.

Real Time PCR—Approximately 50 ng of cDNA was used with TaqMan Universal PCR Master Mix (Applied Biosystems, Warrington, UK) for all real time PCRs. Cycling conditions were as follows: 50 °C for 2 min and 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Details of probes used can be found in Supplemental Table 2.

Nuclear Fractionation—Nuclear fractions of the ES and sorted differentiated cells were obtained with the use of the Nuclear Extract kit (Active Motif, Brussels, Belgium) as described in the manufacturer's instructions. The nuclei were then lysed with 8 M urea. The lysates were homogenized via passage through a 23-gauge needle and then dialyzed in 0.9 M urea.

Mass Spectrometry—Protein samples (100 μg) were prepared as described previously (42), then digested by addition of 5 μg of trypsin at 0.5 μg/μl, and incubated overnight at 37 °C. Peptides were labeled with iTRAQ reagent (Applied Biosystems, Framingham, MA) as described previously (42). Reactions were then checked for complete labeling using mass spectrometry; the four isobarically tagged samples (named 114, 115, 116, and 117 after the reporter ion associated with the tag) were then pooled prior to analysis.

Peptide Fractionation and Mass Spectrometry—Prior to reverse phase LC-MS/MS peptides were fractionated off line using a strong cation exchange chromatography column (10-cm × 2.1-mm PolyLC Polysulfoethyl A column, 5-μm beads, 200-Å pore size (The Nest Group, Southborough, MA) on an LC Packings Ultimate LC system. The gradient was run at 250 μl/min using initially 10 mM KH_2PO_4, 20% ACN, pH 2.7 with KCl concentration increasing from 0 to 250 mM KCl over a 45-min period and then from 250 to 500 mM over 10 min followed by 1 M for 5 min. Five-minute fractions were collected for the first 15 min followed by 1-min fractions. Each fraction was then concentrated. Reverse phase LC-MS/MS was then performed using the procedures described previously (42). For analysis, about 30% of the peptide sample was used on each run with the QSTAR® XL mass spectrometer (Applied Biosystems, Warrington, UK).

Transcriptome Analysis—RNA from GFP[−]Flk1[−], GFP[+]Flk1[−], and GFP[+]Flk1[+] ES cells was prepared in triplicate using the Qiagen RNeasy kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. Transcriptome analysis was undertaken by the Cancer Research UK microarray facility (Paterson Institute, Manchester, UK). Details of the protocols are available upon request. Total RNA was subjected to amplification using the "gene chip eukaryotic small sample target labeling assay version II" developed by Affymetrix, and the gene expression profile was assessed using Mouse 430-2 chips in triplicate.

Data Analysis—Proteome data were searched against a mouse KBMS3.0 protein database from Celera Discovery System (369,691 entries) (Applied Biosystems, Framingham, MA) using ProQUANT version 1.1 (Applied Biosystems) as described previously (42). Default search parameters within ProQUANT were set with trypsin as the digestion agent allowing for one missed cleavage, an MS tolerance of 0.15 Da, and an MS/MS tolerance of 0.1 Da. Fixed modifications included methylmethanethiosulfonate alkylation of cysteine and iTRAQ modification of lysine and N termini. Variable modifications included iTRAQ labeling of tyrosine. Only peptides with a minimum confidence score of 70 were included in the final dataset; this allowed only the highest quality MS/MS spectra to be used for protein identification and therefore relative quantification. Quantitation was performed by comparing iTRAQ reporter ion peak areas where iTRAQ peaks less than 40 counts were not used. Outliers were not removed. A minimum of three peptides were required to identify and relatively quantify a protein. Quantitation was achieved by calculating a weighted average from all matching peptides along with a 95% confidence interval to assess accuracy of protein ratio. Experiments were performed in triplicate, and the ProGROUP (Applied Biosystems) software was used to pool data from all experiments and assign a single accession number for each protein, allowing comparison of replicate datasets. iTRAQ ratios greater than or equal to 1.2 or less than 0.8 found in three experiments with a p value less than 0.05 were flagged as being differentially expressed. Performing the search against a concatenated database containing both forward and reversed sequences (therefore 739,382 entries) allowed estimation of the false discovery level. Transcriptome data were analyzed by submitting each replicate pair individually to GeneChip® Operating Software version 1.2 software (Affymetrix) with default settings with the exception of the target normalization value, which was set to 100. A transcript was determined to be "significantly differentially expressed" if it was called "present" and "increased" or "decreased" in all three replicate arrays.

To compare transcriptome and proteome data, all proteins with a Swiss-Prot, TrEMBL, or RefSeq accession number were selected, and these accession numbers were used to identify the relevant probe set identifiers from the microarray dataset using the NetAffx tool (Affymetrix). The Pathway Analyst software (Stratagene, Amsterdam, The Netherlands) was used to map interaction pathways on protein changes.

Various self-organizing map (SOM) size matrices were tested with a 1 × 3 map being the most appropriate for the generation of distinct cluster groups whose members shared similar expression profiles as cells differentiated. SOMs were trained using the Kohonen algorithm. Two training phases were used. Phase 1 (the ordering phase) consisted of 2000 epochs, an initial learning rate of 0.5 decreasing to 0.1, and a neighborhood size of 1. Phase 2 (the fine tuning phase) consisted of 1000 epochs, a consistent learning rate of 0.1, and a neighborhood size of 0.

Western Blot Analysis—Nuclear fractions and whole cell lysates of ES cells sorted on the characteristics described above were size-fractionated by SDS-PAGE, and proteins were transferred onto nitrocellulose membrane (Hybond-C, Amersham Biosciences) in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, 0.0375% SDS) using an XCell SureLock[TM] (Invitrogen). Membranes were blocked with a 5% (w/v) nonfat milk solution in TTBS (0.05% Tween 20, 20 mM Tris-HCl, pH 7.5, 500 mM NaCl) for 2 h. Immobilized antigen was detected following incubation with primary antibody diluted in 1% (w/v) nonfat milk solution in TTBS overnight. Blots were then washed using TTBS and incubated with the appropriate horseradish peroxidase-conjugated secondary antibody in 5% (w/v) nonfat milk solution in TTBS for 1 h. The blots were then washed using TTBS and developed using ECL (Amersham Biosciences). Antibodies and their sources are shown in Supplemental Table 3.

RESULTS

Selection Based on Flk1 and Brachyury Expression to Generate Cells for Systematic Microarray and Proteomics Analysis of Early Development—When ES cells are cultured in the absence of LIF, in a medium that supports mesodermal development, *Bry*[GFP/+] ES cells display an ability to differentiate. Furthermore the presence of Brachyury at a single allele has been shown to have no adverse effect on development (28). Thus the system offers opportunities for flow cytometric sorting of *Bry*[GFP/+] ES cells based on *GFP* expression enabling the enrichment of cells at specific stages of development. After 2.5 days of differentiation in suspension culture a GFP[−]Flk1[−] population was sorted using a fluorescence-activated cell sorter. Cells of this phenotype have been shown previously to have characteristics of epiblast-like cells. At 3.5 days GFP[+]Flk1[−] (mesoderm-committed cells) and GFP[+]Flk1[+] cells (enriched for hemangioblast-like cells, see Fig. 1*A*) were sorted, and these three populations along with unsorted ES cells were used for microarray or proteomics analysis. Prior to these analyses the developmental stage of the four populations was assessed at both the genomic and
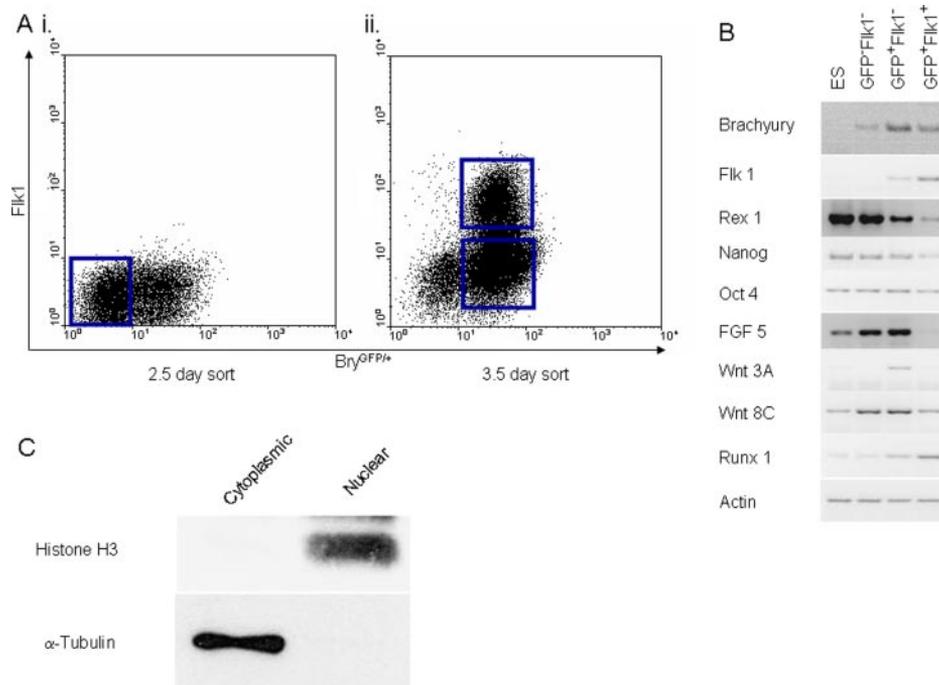
FIG. 1. **Enrichment of ES cells at specific stages of development for nuclear proteomics analysis and comparison with transcriptome data.** *A*, flow cytometric enrichment of $Bry^{GFP/+}$ ES cells based on *GFP* and *Flk1* expression. At least 20 million cells were prepared for each experiment by flow cytometric sorting. *i*, scatter plot of data from time point 2.5 days after LIF withdrawal and cells being placed in suspension culture to promote embryoid body formation; the *blue gate* shows the GFP⁻Flk1⁻ cells. *ii*, cells after 3.5 days without LIF showing the *upper* and *lower blue gates* set for collection of GFP⁺Flk1⁻ and GFP⁺Flk1⁺ cells, respectively. Results are from one of three experiments where cells were taken for proteomics and transcriptomics analysis. *B*, gene analysis of the four populations of cells. RNA was extracted from unsorted $Bry^{GFP/+}$ ES cells and sorted differentiated cells. Semiquantitative RT-PCR was carried out on genes known to be involved in $Bry^{GFP/+}$ ES cell differentiation. *C*, Western blot analysis of known nuclear and cytosolic proteins in fractions prepared from ES cells. $Bry^{GFP/+}$ ES cells were lysed and fractionated, and a nuclear fraction was prepared. Western blot analysis for a cytosolic (α-tubulin) and a nuclear (histone H3) protein was performed to assess nuclear preparation enrichment. Results shown are typical of four fractionation experiments performed.

biological level. Gene expression analysis was performed on RNA extracted from the cell populations (Fig. 1*B*). The expression of genes known to be expressed in these developing cells was determined. The blast cell colony-forming potential of the sorted cell populations was also evaluated, and as anticipated (28), the majority was seen within the GFP⁺Flk1⁺ cells (data not shown). The markers used for flow cytometric enrichment thus offer opportunity for systematic analysis on developmental progression ranging from premesoderm to prehemangioblast to hemangioblast.

Stem cell signatures have been defined at the mRNA level using microarrays (46–48). However, we have shown previously that 54% of proteomic changes in hematopoietic stem cells do not correlate with that seen for specific mRNA species (41). Proteomics technologies require decreased sample complexity to identify lower abundance proteins. We therefore fractionated the $Bry^{GFP/+}$ ES cells, enriching for nuclei. The nuclear fraction contained ~67% of total cellular protein in all four populations. Fig. 1*C* shows that the marker proteins for cytosol and nucleus reside predominantly in the appropriate fractions. The nuclear fractions from the four populations were each trypsinized and labeled with a specific iTRAQ reagent. The samples were then mixed and subjected to

two-dimensional LC-MS/MS. Protein identification and relative quantification was achieved by integrating the peptide data for each specific protein. The protein content per cell was the same in all four populations in all three experiments performed. Eighty-one percent of proteins we identified and quantified had previously been defined with a nuclear localization (49).

The enrichment procedure had the desired effect as peptides derived from low abundance proteins (such as transcription factors, *e.g.* Sox2 (see Fig. 5)) were detected that have not been detected in similar experiments on whole cell lysates. Identification and relative quantification on 2389 proteins (>3 peptides/protein) was achieved (see Supplemental Table 4). Previous studies have shown cutoff values for significant change at the protein level to be ≥1.2 and ≤0.8 (42, 50–54). Unwin *et al.* (42) demonstrated that comparison of biological replicates in one iTRAQ experiment enabled accurate differences in expression to be identified at a ratio of ≥1.2 and ≤0.8 for ≥3 peptides These values have been used as one criterion for proteins that show a change in expression level. To identify differentially expressed proteins a minimum of three peptides were required to identify and relatively quantify a protein. iTRAQ ratios also required a *p* value less than

TABLE I

Correlation of transcriptomics and proteomics comparisons in GFP⁻Flk1⁻, GFP⁺Flk1⁻, and GFP⁺Flk1⁺ Bry^{GFP/+} ES cells

Microarray and iTRAQ data between two cell populations were compared and expressed as up-regulated, down-regulated and not changed. The final group compares the microarray changes seen in the GFP⁻Flk1⁻:GFP⁺Flk1⁻ comparison with the protein changes seen in the GFP⁺Flk1⁻:GFP⁺Flk1⁺ comparison.

	mRNA up	mRNA no change	mRNA down
GFP⁻Flk1⁻:GFP⁺Flk1⁻			
Protein up	8	120	3
Protein unchanged	8	883	18
Protein down	0	13	4
GFP⁺Flk1⁻:GFP⁺Flk1⁺			
Protein up	3	37	0
Protein unchanged	4	854	11
Protein down	0	142	6

	mRNA up	mRNA GFP⁻Flk1⁻:GFP⁺Flk1⁻	mRNA down
GFP⁺Flk1⁻:GFP⁺Flk1⁺			
Protein up	3	35	1
Protein unchanged	12	840	17
Protein down	1	140	7

0.05 (95% confidence limit of proteins considered to change) for the protein to be considered as changed. Of the four populations of cells included in the analysis the undifferentiated the $Bry^{+/GFP}$ ES cell population was the only population that was unsorted. It was therefore a highly heterogenous population of cells that also included apoptotic cells. When comparative analysis was performed on this population it was found to give the greatest variation between the triplicate datasets, and as such these cells were excluded from further analysis. The GFP⁺Flk1⁻ cells compared with GFP⁻Flk1⁻ cells exhibited 8.5% of proteins changing significantly (192 proteins) between these two populations. GFP⁺Flk1⁺ cells showed a 14% change (334 proteins) when compared with GFP⁻Flk1⁻ cells. Furthermore GFP⁺Flk1⁻ cells compared with GFP⁺Flk1⁺ cells showed 9.5% of proteins changed significantly in their level of expression (226 proteins). We next determined how this related to changes in transcription from these genes. Searching against a reversed database allowed calculation of the false discovery rate as 0.15962% at the protein level.
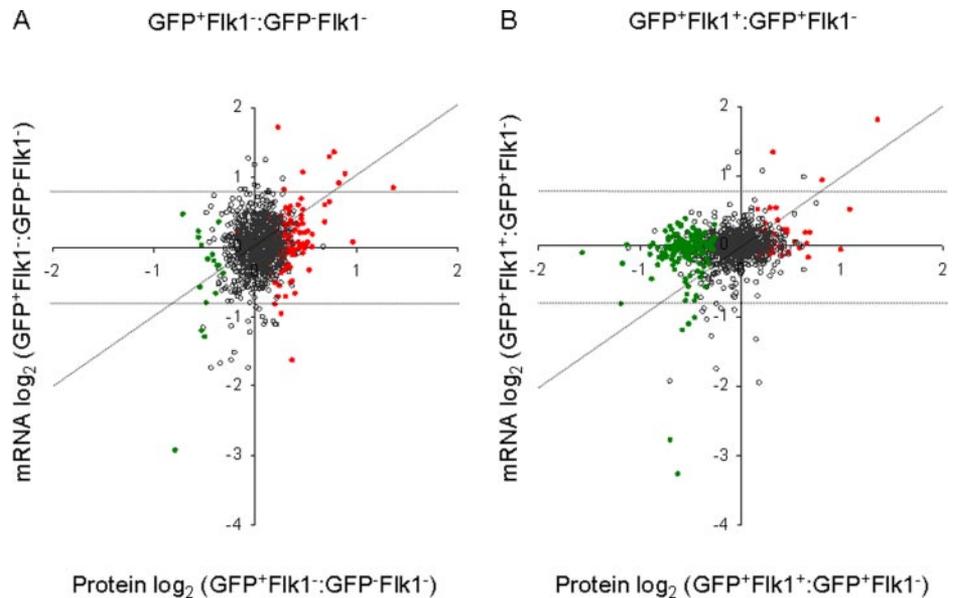
Proteomics and Genomics Analysis of Developing Bry^{GFP/+} ES Cells—Further analysis of these data was performed with the comparison of the protein relative quantification with the microarray data (Supplemental Table 5). We correlated 1057 proteins to probe sets from the microarray analysis. The reason a complete comparison was not made is because of the inability of the NetAffx (Affymetrix) software to match certain protein accession numbers to their associated microarray probe sets with high levels of certainty. We therefore concentrated analysis on data where co-identity of protein and probe set data was assured. Based on an analysis of the matched dataset we found a low correlation between the protein and microarray changes seen for each gene product (Table I). That is, many changes in protein levels could not be ascribed to altered mRNA expression levels. The lack of correlation spanned each pairwise comparison of the three sorted populations. The $Bry^{+/GFP}$ GFP⁺Flk1⁻ cells compared with GFP⁻Flk1⁻ cells showed 120 cases where the protein was significantly increased yet mRNA level was not significantly changed; hence we have identified a substantial number of events that would not have been observed with the use of microarray analysis alone. This trend continued when $Bry^{+/GFP}$ GFP⁺Flk1⁺ cells were compared with GFP⁺Flk1⁻ cells. One hundred and seventy-nine significant protein changes (of a total of 188 protein changes) were identified where the mRNA expression was not altered. When the changes in expression level of the RNA and protein were considered in the GFP⁺Flk1⁻ and GFP⁺Flk1⁺ $Bry^{+/GFP}$ nine of the 188 protein changes correlated to changes in mRNA (Table I), whereas in GFP⁻Flk1⁻ and GFP⁺Flk1⁻ cells, 12 of 148 protein changes were seen at the RNA level as well.

Protein changes from $Bry^{+/GFP}$ GFP⁻Flk1⁻ cells and GFP⁺Flk1⁻ cells primarily involved up-regulation (131 up-regulated and 17 down-regulated proteins). Of these 131 proteins, 48 are subsequently down-regulated as the cells differentiate toward the hemangioblast (GFP⁻Flk1⁻ cells and GFP⁺Flk1⁻ cells). This transitory expression helps to define the mesodermal development of ES cells.

A pairwise comparison of the protein and mRNA ratios between the populations was constructed to determine whether there was any trend in mRNA to protein changes between the two populations (Fig. 2, *A* and *B*). Previous studies have shown that in protein and RNA datasets the majority of genes are not altered in their expression level (41, 42, 54). Thus the data in Fig. 2 and Table I demonstrate that the vast majority of proteins and mRNA identified in this analysis are not altered. Bland-Altman plots were generated to assess the agreement between the protein and RNA ex-

FIG. 2. **The changes in protein expression seen in developing *Bry^GFP/+* ES cells compared with transcriptome changes.** Cells were enriched on the basis of *GFP* and *Flk1* expression, and protein and mRNA data were compared to generate graphs comparing GFP⁻Flk1⁻ with GFP⁺Flk1⁻ (*A*) and GFP⁺Flk1⁻ with GFP⁺Flk1⁺ (*B*). Significant protein increases between two samples in the analysis are denoted by *red data points*, significant decreases are denoted by *green data points*, and unchanging proteins where mRNA and protein data were available are shown as *unfilled circles*. The *lines* represent the levels above and below which ratiometric microarray differences between the two populations shown at the *top* of the figure are considered significant. The *diagonal line* (equation $x = y$) functions as a reference; points falling upon this line correlate in their protein and RNA expression. Data shown are derived from three separate experiments.

A — GFP⁺Flk1⁻:GFP⁻Flk1⁻

B — GFP⁺Flk1⁺:GFP⁺Flk1⁻

(A) y-axis: mRNA log₂ (GFP⁺Flk1⁻:GFP⁻Flk1⁻); x-axis: Protein log₂ (GFP⁺Flk1⁻:GFP⁻Flk1⁻)

(B) y-axis: mRNA log₂ (GFP⁺Flk1⁺:GFP⁺Flk1⁻); x-axis: Protein log₂ (GFP⁺Flk1⁺:GFP⁺Flk1⁻)

pression levels. This test is more appropriate for measuring agreement than a standard correlation coefficient, which measures whether points are linearly related but does not indicate agreement between measurements from different methods. When correlation is used over a large number of data points with a wide range of values, a statistically significant correlation is almost guaranteed, and therefore it is an inappropriate and misleading comparison for high dimensional proteomics and genomics data. The Bland-Altman approach plots the difference against the average of the two measurements, and the regression line of the difference is calculated. Here a difference closer to 0 indicates good agreement between the expression levels. Therefore a *p* value generated from the regression line that is <0.05 would indicate a genuine trend in the differences in expression levels and therefore no agreement. Using this statistically robust approach, no agreement was seen between protein and RNA expression levels (Table II). Next we considered the proteins and mRNA that are changing in these datasets. In Fig. 2 the *lines* represent the levels above and below which ratiometric microarray differences between the two populations are considered significant. Proteins with a significant increase in the specific paired comparison were labeled as *red spots*, and those decreasing were labeled with *green spots*. This graphic method showed no determination could be made on -fold protein change from mRNA change. We can conclude that where changes are seen in protein expression these do not agree significantly with mRNA changes as determined using Bland-Altman testing (Table II). An exception to this finding was seen when the *Bry^+/GFP* GFP⁻Flk1⁻ cells and GFP⁺Flk1⁻ cells were compared, and the proteins and RNA that were identified as up-regulated appear to correlate. Thus mRNA-derived data do not necessarily give an indication of

TABLE II

Agreement of protein and RNA expression in GFP⁻Flk1⁻, GFP⁺Flk1⁻, and GFP⁺Flk1⁺ Bry^GFP/+ ES cells

The measure of agreement was calculated using the regression line of the slope generated by Bland-Altman plots. Here a *p* value <0.05 indicates a statistically significant trend in the differences between protein and RNA expression.

	Protein to RNA regression, R^2	*p* value
GFP⁻Flk1⁻:GFP⁺Flk1⁻		
Protein up	0.254	1.2×10^{-7}
Protein unchanged	0.025	4×10^{-6}
Protein down	0.702	2×10^{-30}
GFP⁺Flk1⁻:GFP⁺Flk1⁺		
Protein up	0.595	2×10^{-26}
Protein unchanged	0.565	8×10^{-165}
Protein down	0.935	9×10^{-18}

the associated change at the protein level in the nucleus (Fig. 2 and Tables I and II). There is neither a quantitative nor a qualitative relationship between changes in protein and mRNA levels for the majority of gene products (see also Supplemental Table 5).

To determine whether specific mRNA changes preceded protein changes (that is to say mRNA changes potentially required more time to influence the proteome) we also compared GFP⁺Flk1⁻:GFP⁺Flk1⁺ *Bry^+/GFP* ES cell proteome changes with GFP⁻Flk1⁻:GFP⁺Flk1⁻ cell transcriptome changes. However, there was only a 5.35% correlation, which infers that there is little evidence for this argument as the association between protein and mRNA changes is as low as in previously described comparisons (Table I). Once again the majority of changes appear to be at the level of translation where many proteins show elevated and diminished expres-

sion with little alteration seen in the mRNA. Taken together these data indicate that there is no strong correlation between expression of nuclear proteins and transcription. Of course, inherent in the process of subcellular fractionation is the fact that some protein changes could be due to redistribution within the cell. Reference to the most complete study on nuclear proteins to date confirms a nuclear localization for most of the proteins in our dataset (49), but subcellular compartmentalization may affect correlation between nuclear protein and mRNA changes. Nonetheless there is a requirement to perform such subcellular fractionation to analyze changes in low abundance proteins (36, 55–59).

Analysis of Gene Ontology on Gene Products Affected at the Protein or mRNA Level—To determine whether (within gene products affected at the protein or mRNA level) there was an enrichment of genes involved in cellular processes that could infer a biologically relevant regulation of ES cell differentiation, protein and RNA data that were identified as changing were categorized by their gene ontology (GO). Fig. 3, *A* and *B*, shows the result of this analysis looking at the two developmental comparisons GFP$^-$Flk1$^-$:GFP$^+$Flk1$^-$ cells and GFP$^+$Flk1$^-$:GFP$^+$Flk1$^+$ cells. The protein and mRNA changes observed are spread between all the major gene ontology categories. Interestingly we see little change at the protein or RNA level in cell proliferation and differentiation gene products; however, there were large protein changes in DNA metabolism and protein biosynthesis. There is more apparent change seen at the RNA than the protein level in the developmental processes, and this includes chromatin packaging and remodeling with proteins such as high mobility group proteins.

A SOM was used to facilitate the identification of groups of proteins with similar expression patterns. A SOM is an unsupervised neural network algorithm that allows for the analysis of large datasets, uncovering patterns in the data and allowing for data visualization by way of a topological map. A SOM begins with random weight vectors, and the learning process is carried out by the modification of these weight vectors toward the data profiles to be mapped. Once training is complete the data are grouped so that different panels in the SOM represent cases whose profiles are similar. Various SOM size matrices were tested, and given that using larger size matrices resulted in empty clusters, a 1 × 3 map was the most appropriate for the generation of distinct cluster groups. The three cluster groups discovered by the SOM algorithm are summarized in Fig. 3, *C*, *D*, and *E*. Fig. 3*F* shows the median centroid for each cluster. Each cluster group has a unique expression pattern with cluster 1 showing an initial increase at the mesoderm stage and then a decrease to hemangioblast. Cluster 2 shows an increase to mesoderm and then an even sharper increase in expression to hemangioblast. Conversely cluster 3 shows the exact opposite with an overall decrease in the proteins belonging to this cluster. Statistical testing by the non-parametric Mann-Whitney *U* test showed the differences

in protein expression to be highly statistically significant, $p \leq 10^{-5}$, between all cluster groups at both the mesoderm and hemangioblast stages. Plotting a scatter plot to show the expression of the individual proteins at the mesoderm and hemangioblast stages also showed clear separation of the three cluster groups identified by the SOM (Fig. 3*G*).

Gene ontology information (Fig. 3, *H*, *I*, and *J*) revealed that the most striking changes are in cluster 1, which contains 16% of proteins with a role in protein biosynthesis, whereas no proteins have this role in cluster 2, and only 3% of proteins have this role in cluster 3. Cluster 2 has a large proportion of proteins involved in transport (9%) and intracellular protein trafficking (15%); however, these proteins are present at much lower percentages in clusters 1 and 2. The majority of protein changes for chromatin packaging and remodeling (6% of protein changes) are grouped into cluster 2. Immunity and defense is also well characterized in cluster 3 (13%) as compared with the other two clusters. It is also of interest that 5% of the proteins are involved in DNA metabolism, and all are associated with cluster 3. This computational approach has identified distinct clusters of proteins from expression profiling data that enable the identification of proteins sharing similar characteristics and also show a possible relation with biological process.

Pathway Analysis of Proteins Changing in Expression Levels during Differentiation—Proteins identified with a pivotal role in biological processes can be identified by a consideration of the connection between gene products using pathway analysis. Proteins associated to their respective RNA probe set showing altered expression in differentiating ES cells were therefore systematically assessed. This allowed identification of connections between the proteins that have altered expression levels in differentiating *Bry$^{+/GFP}$* ES cells (Fig. 4, *A–C*). In the three comparisons made, GFP$^+$Flk1$^+$:GFP$^-$Flk1$^-$, GFP$^+$Flk1$^-$:GFP$^-$Flk1$^-$, and GFP$^+$Flk1$^+$:GFP$^+$Flk1$^-$, a number of proteins appear to be involved in the regulation of several gene products undergoing altered expression. RBL2 (also known as p130) has a central role in gene regulation. RBL2 is a member of the retinoblastoma (Rb) family of proteins and has been identified previously as being important in stem cell differentiation in several systems (60–62). RBL2 has been shown to have an antiproliferative activity involving the histone deacetylase (HDAC) pathway (61). RBL2 and HDAC1 have been shown to function together to inhibit transcription of the E2F-dependent promoter of cyclin A (63). E2F1 has also been included in the pathway analysis and is demonstrated to be associated with proteins whose expression level increases as the cells differentiate from GFP$^-$Flk1$^-$ to GFP$^+$Flk1$^-$. E2F1 is a transcription factor involved in the regulation of cell cycle mediators, *e.g.* cyclin E (64). High mobility group protein A2 (HMGA2) has been demonstrated to be involved in the regulation of key transcription factors such as E2F1 through its regulation of Rb protein phosphorylation (65–68). In our study HMGA2 increases at the mRNA and protein level (see Fig. 5).
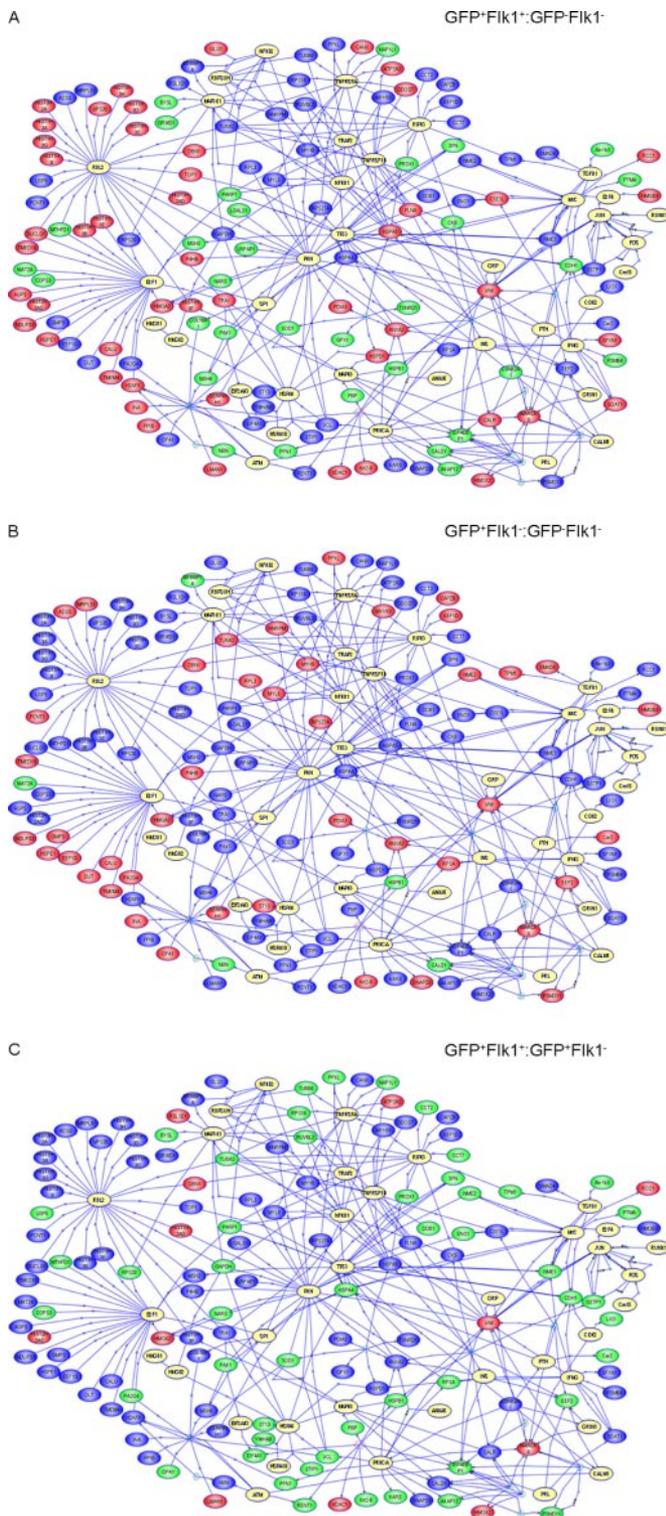
FIG. 3. **Gene ontology of proteins whose levels change as *Bry^{GFP/+}* ES cells differentiate.** The gene ontology of all the proteins showing significant changes in expression levels between the three populations was analyzed by placing the proteins into 25 GO categories. *A* and *B* show a comparison of two *Bry^{GFP/+}* ES cell populations at specific developmental stages. The mRNA changes (*dark gray histograms*) show the percentage of gene products within the dataset exhibiting a change. Values for each GO group were calculated as a percentage of total mRNA changes seen. The percentages of protein changes seen for each GO category are shown as *light gray histograms*. The changes are calculated with relation to total protein changes. Using a 1 × 3 map SOM size matrix proteins whose levels change were clustered into three groups (*C*, *D*, and *E*). *F* shows the median centroid for each cluster. *G*, a scatter plot to show the expression of individual proteins at the mesoderm and hemangioblast stages. *H*, *I*, and *J*, gene ontology information from the three clusters

FIG. 4. **Pathway analysis of proteomic changes in *Bry*^GFP/+ ES cells as they undergo differentiation.** Pathway analysis of the protein changes seen in the three comparisons made between *Bry*^GFP/+ ES cell populations at specific developmental stages is shown. See figure for the comparisons made. *Blue* represents proteins identified that showed an unchanged expression level. *Red* represents a ratiometric value of 1.2 or more, an increase in protein expression, whereas *green* represents a value of less than 0.8 and hence a

This analysis can therefore be used to identify novel proteins and pathways that could be involved in regulating stem cell differentiation. Other proteins that seem to be involved in the regulation of a number of genes include TNFRSFA, also known as tumor necrosis factor receptor superfamily, member 1A. Tumor necrosis factor α has been shown to inhibit hematopoietic stem cell expansion and renewal (69).

Confirmation of Isobaric Tag/Tandem Mass Spectrometry Data on Protein Expression—To confirm the value of the iTRAQ approach for relative quantification in differentiating *Bry*^GFP/+ ES cells we performed Western blot analysis on a number of proteins and coupled this to the real time PCR analysis and microarray data (Fig. 5). Previously UTF1, Sox2, and Oct4 have been demonstrated to be involved in maintenance of pluripotency in ES cells (13–16, 70). As confirmation of our protein and mRNA dataset the kinetics of Oct4, Sox2, and UTF1 were analyzed. It can be seen in Fig. 5*A* as the cells differentiate toward GFP^+Flk1^+ ES cells these gene products are down-regulated at the mRNA and protein level as anticipated. The Western blot and real time PCR data also supported the microarray and the iTRAQ data, respectively, on nuclear proteins involved in chromatin packaging and remodeling or DNA metabolism such as HDAC2, HMGA2, and DNA methyltransferase 1 (DNMT1) (Fig. 5*B*). The Western blots and real time PCR experiments performed here confirm the data on these proteins and RNAs obtained from the microarray and iTRAQ method. The iTRAQ method for relative quantification of 2389 nuclear proteins was successful with the material required to run just a few Western blots. In Fig. 2, *A–C*, and Table I we demonstrated that there was limited correlation between protein and mRNA expression. We identified a number of differences at the mRNA level that did not change at the protein level. An example of this is for DNMT1: the Western blot data and the iTRAQ data were in agreement but did not reflect the situation seen with real time PCR analysis of mRNA levels or microarray data. Systematic proteomics analysis correlated to Western blot analysis in the proteins highlighted here (see Fig. 5). Thus the isobaric tag data obtained can be compared with microarray data (see Fig. 5), which in turn are validated by real time RT-PCR data. This represents a high penetrance approach to understanding stem cell development.

DISCUSSION

The process of embryonic cell development to mesoderm-committed cells and hemangioblast can be studied using the *Bry*^GFP/+ ES cells (28). As they offer sufficient biological material for systematic analysis, we have taken advantage of the possibility for enriching cells at a specific stage of development to undertake mRNA and proteomics analyses. The latter analysis was considered essential as there is evidence of a

decrease in expression levels. *Yellow* represents proteins that were not found in our proteomics analysis but have previously been associated with proteins that were identified.
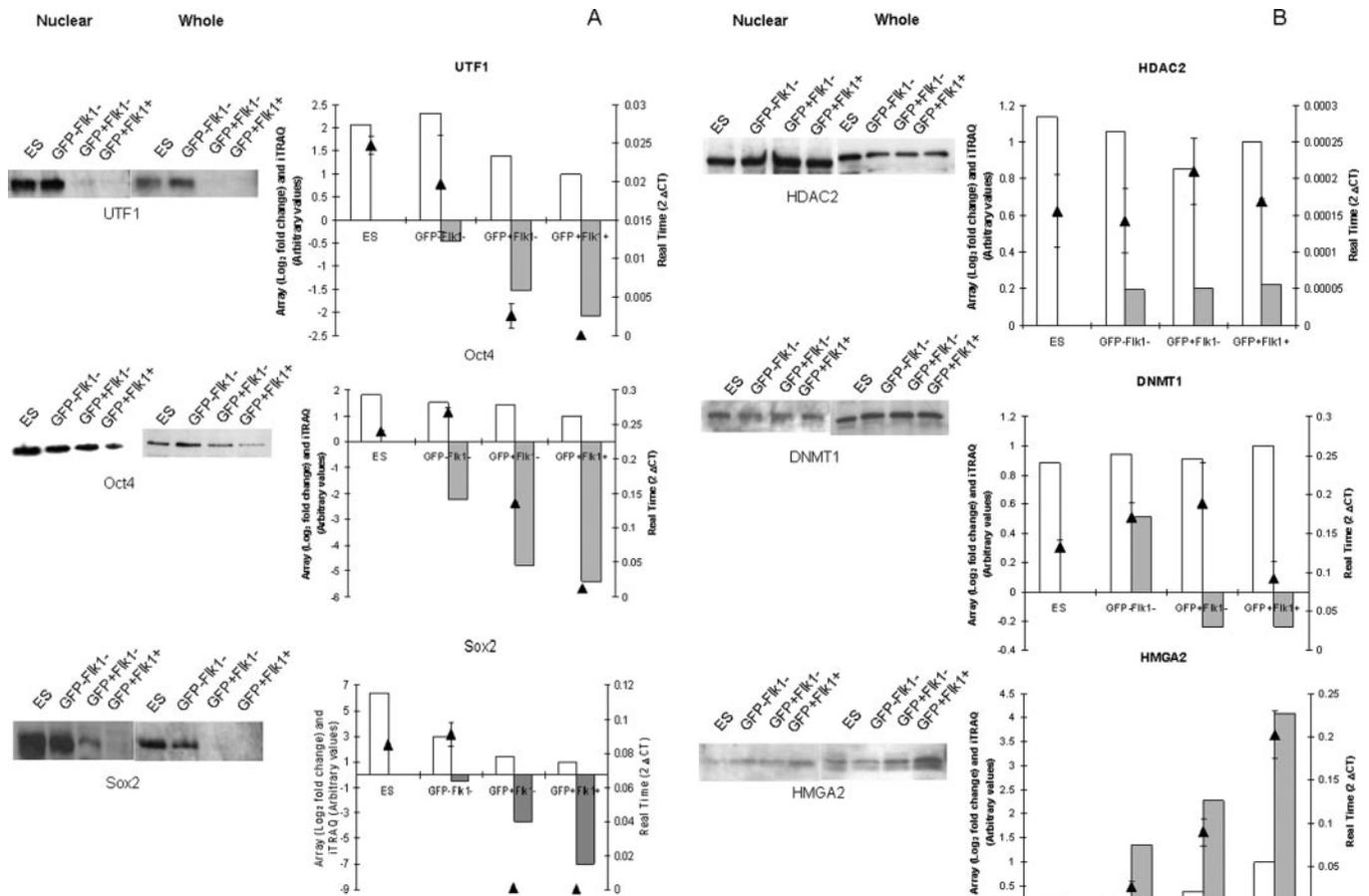
FIG. 5. **A comparison of relative expression levels of proteins and mRNA validated with the use of real time PCR and Western blot.** *A*, Oct4, UTF1, and Sox2, known regulators of pluripotency, were identified in our proteomics and microarray analysis and validated with the use of Western blot and real time PCR, respectively. *B*, validation of the protein and mRNA expression levels of HMGA2, HDAC2, and DNMT1. Western blots were performed with both nuclear and whole cell preparations. The iTRAQ data for each protein is shown (*white histogram*) with data from real time PCR analysis (*line*) and microarray (*gray histogram*). *Error bars* shown are from three observations and represent the S.E.

disconnection between mRNA and protein expression levels for a variety of gene products in enriched hematopoietic stem cells as they undergo development (41). Here we confirm that this observation extends to *Bry*^GFP/+^ ES cells differentiating toward a mesodermal lineage cell fate. Although proteomics studies have been performed on ES cells using two-dimensional gel electrophoresis (71–73) we have developed a novel form of analysis with increased penetration and quantitative accuracy (42) that is now applied to ES cells. The benefits of using a quantitative multidimensional LC-MS/MS approach include a greater ability to analyze hydrophobic, large, and basic proteins, the latter of importance with respect to nuclear proteins such as histones. Furthermore in our hands multidimensional LC-MS/MS incorporating iTRAQ methodologies offer greater proteomic penetration than two-dimensional gel techniques (74, 75).

Our data show that we have identified and relatively quantified 2389 proteins of differentiating *Bry*^GFP/+^ ES cells. This is one of the largest relatively quantified mammalian proteomics datasets published to date. The preparation of nuclei for

proteomics analysis was a requirement to identify changes in transcriptional mechanisms associated with differentiation. It is also widely accepted that prefractionation is required for systematic analysis of low abundance proteins (57–59). Proof of this principle is demonstrated in this report with the identification and relative quantification of low abundance transcription factors, *e.g.* Sox2, Oct4, UTF1, and E2F4. The majority of proteins quantified have been shown to reside (at least in part) in the nucleus (49). This includes proteins such as actin and cofilin, which have been regarded as non-nuclear proteins until recently (76, 77).

Having enriched ES cell populations at specific stages of development and compared proteomic changes to microarray analyses we observe that there is poor statistical agreement between the two. This is supported by previous studies (36–40). Interestingly Unwin *et al.* (41) described an mRNA and proteome comparison of murine hematopoietic stem cells and demonstrated a lack of correlation in 54.7% of the identified genes. In our study the comparison of GFP⁻Flk1⁻ cells with GFP⁺Flk1⁻ cells demonstrated that 89.9% of the pro-

teins that showed a significant change were shown as no change in the microarray analysis. The GFP$^+$Flk1$^-$ cells compared with the GFP$^+$Flk1$^+$ Bry$^{+/GFP}$ cells also demonstrated a pattern of more protein changes than RNA changes (188 compared with 24), and more interesting still 95% of protein changes were not associated with mRNA changes (see Table I). A statistical agreement between protein and mRNA changes would indicate regulation of these genes at the level of chromatin remodeling and transcription (78). However, as discussed, the majority of the significant changes seen at the mRNA or protein levels do not agree. Regulation of these genes must be at the level of translation and/or the relative degradation rates of the mRNA and protein. Alternatively there is a possibility that our results are explained by a temporal lag between changes in mRNA levels and changes in protein levels. We have addressed this issue by comparing early mRNA changes in GFP$^-$Flk1$^-$:GFP$^+$Flk1$^-$ES cells with later stage changes in the proteome (GFP$^+$Flk1$^-$: GFP$^+$Flk1$^+$), yet there are still few correlations between mRNA and protein modulations. Where small incremental changes in development are being compared there will be low correlation between protein and mRNA changes. Stem cells have been shown to produce mRNA for a large number of genes, and this decreases with differentiation (79). One hypothesis is that the mRNA species generated in stem cells are not directly linked to production of protein. They are produced to offer multiple options for development. An extension of this is that as cells develop increased and decreased expression of mRNA species offers differential roads to develop. This suggestion supports the need for systematic analysis at the protein level as well as a microarray-based experiment. Nonetheless the extensive mRNA analysis of ES cells to date has of course yielded data on pathways for self-renewal and differentiation (80, 81). Networks of transcription factor expression and transcription factor protein interaction can be seen to govern differentiation. Our data indicate an increased level of complexity to this system; whereas transcription factor complexes bind to promoter regions to initiate mRNA production, the direct relation between these critical events and their downstream effects are now less clear. Modulation of transcription does not directly govern the levels of many nuclear proteins. This offers a further layer of complexity to the governance of normal differentiation processes. The biological and analytical impact of this piece of work is that it demonstrates that changes at the RNA level cannot be used to assume protein expression levels. Recently Doss *et al.* (82) have shown that an embryonic stem cell line that was differentiated and analyzed using microarray demonstrated 479 genes up-regulated and 193 genes down-regulated. However, due to the lack of agreement seen between protein and RNA expression levels in differentiating stem cells these changes may not be seen at the protein level. Post-transcriptional regulation by micro-RNA is one mechanism that can explain the data reported here (83). Another is the compart-

mentalization of (well defined) nuclear proteins to the cytosol, an unlikely scenario for so many proteins.

We confirmed selected iTRAQ-derived information on modulation of protein levels using Western blot analysis. This gives a high level of confidence that isobaric tag-based mass spectrometry can permit the sensitive relative quantification data acquisition that permits the above conclusions. DNMT1 associates with HDAC proteins and cooperates with retinoblastoma protein to repress transcription from promoters containing E2F-binding sites (84–86). DNA methylation and histone deacetylation are key in the regulation of transcription, yet these two proteins show invariant expression in early ES cell development (Fig. 5). In the case of DNMT1, however, the mRNA level is shown to decrease during differentiation (see Fig. 5). UTF1 is a transcriptional co-activator that has been found to be expressed mainly in pluripotent ES cells (87). The regulatory elements governing UTF1 gene expression include octamer binding transcription factor Oct4. Our data clearly show that Oct4 protein expression decreases as UTF1 expression decreases (Fig. 5). Perhaps surprising is the fact that the Oct4 protein only decreases to a degree before UTF1 becomes undetectable by Western blot. These are examples of the opportunities arising from isobaric tag datasets for network building on gene expression/protein regulation.

We developed an interaction network of proteins identified as changing in expression during ES cell development (Fig. 4). We identified HMGA2 as a protein changing in expression. This protein has a linkage to E2F1, E2F4, and SP1 as well as to HDAC1 and HMGB2. HMG proteins are transcription factors that are expressed during embryonic development (88, 89) and are involved in a network of protein-protein and protein-DNA interactions resulting in the formation of enhanceosomes at promoters (90, 91). *In vitro* studies have demonstrated that the HMGA2 protein was able to efficiently bend DNA and form DNA circles, a role of potential significance in V(D)J recombination (92). These studies suggest a role in facilitating cooperative interactions between cis-acting proteins by promoting DNA flexibility. The HMGA2 gene is commonly overexpressed in neoplastic cells and is frequently amplified in adenomas, apparently affecting tumor formation via the E2F1 protein (93). Knock-out mouse studies of the HMGA2 gene led to a pygmy phenotype (89). Bmi1 is a target of E2F1 and has a pivotal role in embryonic development and self-renewal in embryonic and hematopoietic stem cells (94, 95). Thus the proteomics data we have acquired offers opportunities to build hypotheses on processes for self-renewal regulation in primitive cells. The HMGB2 protein (identified by iTRAQ and microarray; no change in expression levels) is known to interact with HMGA2 and Oct4 (96). The related HMGB3 protein (seen to increase when the cells differentiate from epiblast to mesoderm at the protein and mRNA level by iTRAQ and microarray) regulates the generation of myeloid and lymphoid progenitor cells in the hematopoietic system. It also regulates the self-renewal of hematopoietic stem cells

(97, 98). These data exemplify the opportunities that deep proteomics mining of ES cells undergoing differentiation offers to systematically analyzing gene product function in development.

It can be argued that the use of the system we have described is a rapid screening system to precede more focused hypothesis-driven research. Thus the development of techniques to study the translational regulation of protein levels decreases the reliance on proteomics confirmation of mRNA-based analyses and offers opportunities to develop new quantitative biology approaches to the understanding of stem cell self-renewal or differentiation with a greater probability of discovering proteins that regulate development. Protein interactions can similarly be mapped in ES cells for systematic discovery of regulatory pathways (99). Here we have presented validation of a discovery proteomics approach to stem cell biology.

REFERENCES

1. Wilkinson, D. G., Bhatt, S., and Herrmann, B. G. (1990) Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343,** 657–659
2. Herrmann, B. G. (1991) Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. *Development* **113,** 913–917
3. Kispert, A., and Herrmann, B. G. (1994) Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev. Biol.* **161,** 179–193
4. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J., and Keller, G. (2004) Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432,** 625–630
5. Evans, M. J., and Kaufman, M. H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292,** 154–156
6. Keller, G. M. (1995) In vitro differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* **7,** 862–869
7. Smith, A. G. (2001) Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* **17,** 435–462
8. Keller, G., Kennedy, M., Papayannopoulou, T., and Wiles, M. V. (1993) Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell. Biol.* **13,** 473–486
9. Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999) Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126,** 5073–5084
10. Vittet, D., Prandini, M. H., Berthier, R., Schweitzer, A., Martin-Sisteron, H., Uzan, G., and Dejana, E. (1996) Embryonic stem cells differentiate *in vitro* to endothelial cells through successive maturation steps. *Blood* **88,** 3424–3431
11. Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C., and Keller, G. (1998) A common precursor for hematopoietic and endothelial cells. *Development* **125,** 725–732
12. Ogawa, M., Fraser, S., Fujimoto, T., Endoh, M., Nishikawa, S., and Nishikawa, S. I. (2001) Origin of hematopoietic progenitors during embryogenesis. *Int. Rev. Immunol.* **20,** 21–44
13. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113,** 643–655
14. Chambers, I., and Smith, A. (2004) Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23,** 7150–7160
15. Niwa, H. (2001) Molecular mechanism to maintain stem cell renewal of ES cells. *Cell Struct. Funct.* **26,** 137–148
16. Niwa, H., Masui, S., Chambers, I., Smith, A. G., and Miyazaki, J. (2002) Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. *Mol. Cell. Biol.* **22,** 1526–1536
17. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113,** 631–642
18. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95,** 379–391
19. Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12,** 2048–2060
20. Niwa, H., Miyazaki, J., and Smith, A. G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24,** 372–376
21. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., and Young, R. A. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122,** 947–956
22. D'Souza, S. L., Elefanty, A. G., and Keller, G. (2005) SCL/Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development. *Blood* **105,** 3862–3870
23. Faloon, P., Arentson, E., Kazarov, A., Deng, C. X., Porcher, C., Orkin, S., and Choi, K. (2000) Basic fibroblast growth factor positively regulates hematopoietic development. *Development* **127,** 1931–1941
24. Lacaud, G., Gore, L., Kennedy, M., Kouskoff, V., Kingsley, P., Hogan, C., Carlsson, L., Speck, N., Palis, J., and Keller, G. (2002) Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development *in vitro*. *Blood* **100,** 458–466
25. Robertson, S. M., Kennedy, M., Shannon, J. M., and Keller, G. (2000) A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1. *Development* **127,** 2447–2459
26. Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F. W., and Orkin, S. H. (1996) The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **86,** 47–57
27. Kispert, A., and Herrmann, B. G. (1993) The Brachyury gene encodes a novel DNA binding protein. *EMBO J.* **12,** 3211–3220
28. Fehling, H. J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., and Kouskoff, V. (2003) Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* **130,** 4217–4227
29. Kinder, S. J., Tsang, T. E., Quinlan, G. A., Hadjantonakis, A. K., Nagy, A., and Tam, P. P. (1999) The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* **126,** 4691–4701
30. Shalaby, F., Ho, J., Stanford, W. L., Fischer, K. D., Schuh, A. C., Schwartz, L., Bernstein, A., and Rossant, J. (1997) A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis. *Cell* **89,** 981–990
31. Shalaby, F., Rossant, J., Yamaguchi, T. P., Gertsenstein, M., Wu, X. F., Breitman, M. L., and Schuh, A. C. (1995) Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* **376,** 62–66
32. Angers, S., Thorpe, C. J., Biechele, T. L., Goldenberg, S. J., Zheng, N., MacCoss, M. J., and Moon, R. T. (2006) The KLHL12-Cullin-3 ubiquitin ligase negatively regulates the Wnt-beta-catenin pathway by targeting Dishevelled for degradation. *Nat. Cell Biol.* **8,** 348–357
33. Hernandez-Hernandez, A., Ray, P., Litos, G., Ciro, M., Ottolenghi, S., Beug, H., and Boyes, J. (2006) Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *EMBO J.* **25,** 3264–3274

34. Ji, Z., Degerny, C., Vintonenko, N., Deheuninck, J., Foveau, B., Leroy, C., Coll, J., Tulasne, D., Baert, J. L., and Fafeur, V. (2007) Regulation of the Ets-1 transcription factor by sumoylation and ubiquitinylation. *Oncogene* **26,** 395–406

35. Zhang, Q., Zhang, L., Wang, B., Ou, C. Y., Chien, C. T., and Jiang, J. (2006) A hedgehog-induced BTB protein modulates hedgehog signaling by degrading Ci/Gli transcription factor. *Dev. Cell* **10,** 719–729

36. Anderson, L., and Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18,** 533–537

37. Bro, C., Regenberg, B., Lagniel, G., Labarre, J., Montero-Lomeli, M., and Nielsen, J. (2003) Transcriptional, proteomic, and metabolic responses to lithium in galactose-grown yeast cells. *J. Biol. Chem.* **278,** 32141–32149

38. Grolleau, A., Bowman, J., Pradet-Balade, B., Puravs, E., Hanash, S., Garcia-Sanz, J. A., and Beretta, L. (2002) Global and specific translational control by rapamycin in T cells uncovered by microarrays and proteomics. *J. Biol. Chem.* **277,** 22175–22184

39. Harris, M. N., Ozpolat, B., Abdi, F., Gu, S., Legler, A., Mawuenyega, K. G., Tirado-Gomez, M., Lopez-Berestein, G., and Chen, X. (2004) Comparative proteomic analysis of all-trans-retinoic acid treatment reveals systematic posttranscriptional control mechanisms in acute promyelocytic leukemia. *Blood* **104,** 1314–1323

40. Maziarz, M., Chung, C., Drucker, D. J., and Emili, A. (2005) Integrating global proteomic and genomic expression profiles generated from islet alpha cells: opportunities and challenges to deriving reliable biological inferences. *Mol. Cell. Proteomics* **4,** 458–474

41. Unwin, R. D., Smith, D. L., Blinco, D., Wilson, C. L., Miller, C. J., Evans, C. A., Jaworska, E., Baldwin, S. A., Barnes, K., Pierce, A., Spooncer, E., and Whetton, A. D. (2006) Quantitative proteomics reveals posttranslational control as a regulatory factor in primary hematopoietic stem cells. *Blood* **107,** 4687–4694

42. Unwin, R. D., Pierce, A., Watson, R. B., Sternberg, D. W., and Whetton, A. D. (2005) Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol. Cell. Proteomics* **4,** 924–935

43. Kubo, A., Shinozaki, K., Shannon, J. M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H. J., and Keller, G. (2004) Development of definitive endoderm from embryonic stem cells in culture. *Development* **131,** 1651–1662

44. Lowe, D. G., Nunes, W., Bombara, M., McCabe, S., Ranges, G. E., Henzel, W., Tomida, M., Yamamoto-Yamaguchi, Y., Hozumi, M., and Goeddel, D. V. (1989) Genomic cloning and heterologous expression of human differentiation-stimulating factor. *DNA* **8,** 351–359

45. Schmelzer, C. H., Harris, R. J., Butler, D., Yedinak, C. M., Wagner, K. L., and Burton, L. E. (1993) Glycosylation pattern and disulfide assignments of recombinant human differentiation-stimulating factor. *Arch. Biochem. Biophys.* **302,** 484–489

46. Furusawa, T., Ikeda, M., Inoue, F., Ohkoshi, K., Hamano, T., and Tokunaga, T. (2006) Gene expression profiling of mouse embryonic stem cell subpopulations. *Biol. Reprod.* **75,** 555–561

47. Hailesellasse Sene, K., Porter, C. J., Palidwor, G., Perez-Iratxeta, C., Muro, E. M., Campbell, P. A., Rudnicki, M. A., Andrade-Navarro, M. A., Lindsley, R. C., Gill, J. G., Kyba, M., Murphy, T. L., Murphy, K. M., Furusawa, T., Ikeda, M., Inoue, F., Ohkoshi, K., Hamano, T., and Tokunaga, T. (2007) Gene function in early mouse embryonic stem cell differentiation. *BMC Genomics* **8,** 85

48. Lindsley, R. C., Gill, J. G., Kyba, M., Murphy, T. L., Murphy, K. M., Furusawa, T., Ikeda, M., Inoue, F., Ohkoshi, K., Hamano, T., and Tokunaga, T. (2006) Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development* **133,** 3787–3796

49. Kislinger, T., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M. S., Gramolini, A. O., Morris, Q., Hallett, M. T., Rossant, J., Hughes, T. R., Frey, B., and Emili, A. (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* **125,** 173–186

50. Duthie, K. A., Osborne, L. C., Foster, L. J., and Abraham, N. (2007) Proteomics Analysis of Interleukin (IL)-7-induced Signaling Effectors Shows Selective Changes in IL-7Rα449F Knock-in T Cell Progenitors. *Mol. Cell. Proteomics* **6,** 1700–1710

51. Guo, Y., Singleton, P. A., Rowshan, A., Gucek, M., Cole, R. N., Graham, D. R., Van Eyk, J. E., and Garcia, J. G. (2007) Quantitative proteomics analysis of human endothelial cell membrane rafts: evidence of MARCKS and MRP regulation in the sphingosine 1-phosphate-induced barrier enhancement. *Mol. Cell. Proteomics* **6,** 689–696

52. Islinger, M., Luers, G. H., Li, K. W., Loos, M., and Volkl, A. (2007) Rat liver peroxisomes after fibrate treatment. A survey using quantitative mass spectrometry. *J. Biol. Chem.* **282,** 23055–23069

53. Jones, A. M., Bennett, M. H., Mansfield, J. W., and Grant, M. (2006) Analysis of the defence phosphoproteome of Arabidopsis thaliana using differential mass tagging. *Proteomics* **6,** 4155–4165

54. Seth, G., Philp, R. J., Lau, A., Jiun, K. Y., Yap, M., and Hu, W. S. (2007) Molecular portrait of high productivity in recombinant NS0 cells. *Biotechnol. Bioeng.* **97,** 933–951

55. Unwin, R. D., Evans, C. A., and Whetton, A. D. (2006) Relative quantification in proteomics: new approaches for biochemistry. *Trends Biochem. Sci.* **31,** 473–484

56. Yates, J. R., III, Gilchrist, A., Howell, K. E., and Bergeron, J. J. (2005) Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell. Biol.* **6,** 702–714

57. Ho, E., Hayen, A., and Wilkins, M. R. (2006) Characterisation of organellar proteomes: A guide to subcellular proteomic fractionation and analysis. *Proteomics* **6,** 5746–5757

58. Klose, J. (1999) Fractionated extraction of total tissue proteins from mouse and human for 2-D electrophoresis. *Methods Mol. Biol.* **112,** 67–85

59. Park, Y. M., Kim, J. Y., Kwon, K. H., Lee, S. K., Kim, Y. H., Kim, S. Y., Park, G. W., Lee, J. H., Lee, B., and Yoo, J. S. (2006) Profiling human brain proteome by multi-dimensional separations coupled with MS. *Proteomics* **6,** 4978–4986

60. Garriga, J., Limon, A., Mayol, X., Rane, S. G., Albrecht, J. H., Reddy, E. P., Andres, V., and Grana, X. (1998) Differential regulation of the retinoblastoma family of proteins during cell proliferation and differentiation. *Biochem. J.* **333,** 645–654

61. Jori, F. P., Melone, M. A., Napolitano, M. A., Cipollaro, M., Cascino, A., Giordano, A., and Galderisi, U. (2005) RB and RB2/p130 genes demonstrate both specific and overlapping functions during the early steps of *in vitro* neural differentiation of marrow stromal stem cells. *Cell Death Differ.* **12,** 65–77

62. Lipinski, M. M., and Jacks, T. (1999) The retinoblastoma gene family in differentiation and development. *Oncogene* **18,** 7873–7882

63. Stiegler, P., De Luca, A., Bagella, L., and Giordano, A. (1998) The COOH-terminal region of pRb2/p130 binds to histone deacetylase 1 (HDAC1), enhancing transcriptional repression of the E2F-dependent cyclin A promoter. *Cancer Res.* **58,** 5049–5052

64. Sherr, C. J., and Roberts, J. M. (1999) CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev.* **13,** 1501–1512

65. Fedele, M., Pierantoni, G. M., Visone, R., and Fusco, A. (2006) E2F1 activation is responsible for pituitary adenomas induced by HMGA2 gene overexpression. *Cell Div.* **1,** 17

66. Fedele, M., Visone, R., De Martino, I., Troncone, G., Palmieri, D., Battista, S., Ciarmiello, A., Pallante, P., Arra, C., Melillo, R. M., Helin, K., Croce, C. M., and Fusco, A. (2006) HMGA2 induces pituitary tumorigenesis by enhancing E2F1 activity. *Cancer Cell* **9,** 459–471

67. Muller, H., Bracken, A. P., Vernell, R., Moroni, M. C., Christians, F., Grassilli, E., Prosperini, E., Vigo, E., Oliner, J. D., and Helin, K. (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev.* **15,** 267–285

68. Vigo, E., Muller, H., Prosperini, E., Hateboer, G., Cartwright, P., Moroni, M. C., and Helin, K. (1999) CDC25A phosphatase is a target of E2F and is required for efficient E2F-induced S phase. *Mol. Cell. Biol.* **19,** 6379–6395

69. Zhang, X., Sejas, D. P., Qiu, Y., Williams, D. A., and Pang, Q. (2007) Inflammatory ROS promote and cooperate with the Fanconi anemia mutation for hematopoietic senescence. *J. Cell Sci.* **120,** 1572–1583

70. Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999) The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell. Biol.* **19,** 5453–5465

71. Kurisaki, A., Hamazaki, T. S., Okabayashi, K., Iida, T., Nishine, T., Chonan, R., Kido, H., Tsunasawa, S., Nishimura, O., Asashima, M., and Sugino, H. (2005) Chromatin-related proteins in pluripotent mouse embryonic stem cells are downregulated after removal of leukemia inhibitory factor. *Biochem. Biophys. Res. Commun.* **335,** 667–675

72. Wang, D., and Gao, L. (2005) Proteomic analysis of neural differentiation of mouse embryonic stem cells. *Proteomics* **5,** 4414–4426

73. Yin, X., Mayr, M., Xiao, Q., Mayr, U., Tarelli, E., Wait, R., Wang, W., and Xu, Q. (2005) Proteomic dataset of Sca-1+ progenitor cells. *Proteomics* **5,** 4533–4545

74. Evans, C. A., Tonge, R., Blinco, D., Pierce, A., Shaw, J., Lu, Y., Hamzah, H. G., Gray, A., Downes, C. P., Gaskell, S. J., Spooncer, E., and Whetton, A. D. (2004) Comparative proteomics of primitive hematopoietic cell populations reveals differences in expression of proteins regulating motility. *Blood* **103,** 3751–3759

75. Unwin, R. D., and Whetton, A. D. (2006) Systematic proteome and transcriptome analysis of stem cell populations. *Cell Cycle* **5,** 1587–1591

76. Blessing, C. A., Ugrinova, G. T., and Goodson, H. V. (2004) Actin and ARPs: action in the nucleus. *Trends Cell Biol.* **14,** 435–442

77. Rando, O. J., Zhao, K., and Crabtree, G. R. (2000) Searching for a function for nuclear actin. *Trends Cell Biol.* **10,** 92–97

78. Claycomb, J. M., and Orr-Weaver, T. L. (2005) Developmental gene amplification: insights into DNA replication and gene expression. *Trends Genet.* **21,** 149–162

79. Bruno, L., Hoffmann, R., McBlane, F., Brown, J., Gupta, R., Joshi, C., Pearson, S., Seidl, T., Heyworth, C., and Enver, T. (2004) Molecular signatures of self-renewal, differentiation, and lineage choice in multipotential hemopoietic progenitor cells *in vitro. Mol. Cell. Biol.* **24,** 741–756

80. Pritsker, M., Ford, N. R., Jenq, H. T., and Lemischka, I. R. (2006) Genome-wide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 6946–6951

81. Sherwood, R. I., Jitianu, C., Cleaver, O., Shaywitz, D. A., Lamenzo, J. O., Chen, A. E., Golub, T. R., and Melton, D. A. (2007) Prospective isolation and global gene expression analysis of definitive and visceral endoderm. *Dev. Biol.* **304,** 541–555

82. Doss, M. X., Chen, S., Winkler, J., Hippler-Altenburg, R., Odenthal, M., Wickenhauser, C., Balaraman, S., Schulz, H., Hummel, O., Hubner, N., Ghosh-Choudhury, N., Sotiriadou, I., Hescheler, J., and Sachinidis, A. (2007) Transcriptomic and phenotypic analysis of murine embryonic stem cell derived BMP2+ lineage cells: an insight into mesodermal patterning. *Genome Biol.* **8,** R184

83. Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006) MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. *Cold Spring Harb. Symp. Quant. Biol.* **71,** 523–530

84. Kimura, H., Nakamura, T., Ogawa, T., Tanaka, S., and Shiota, K. (2003) Transcription of mouse DNA methyltransferase 1 (Dnmt1) is regulated by both E2F-Rb-HDAC-dependent and -independent pathways. *Nucleic Acids Res.* **31,** 3101–3113

85. Morrison, A. J., Sardet, C., and Herrera, R. E. (2002) Retinoblastoma protein transcriptional repression through histone deacetylation of a single nucleosome. *Mol. Cell. Biol.* **22,** 856–865

86. Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L., and Wolffe, A. P. (2000) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.* **25,** 338–342

87. 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 extraembryonic cells. *EMBO J.* **17,** 2019–2032

88. Cerignoli, F., Ambrosi, C., Mellone, M., Assimi, I., di Marcotullio, L., Gulino, A., and Giannini, G. (2004) HMGA molecules in neuroblastic tumors. *Ann. N. Y. Acad. Sci.* **1028,** 122–132

89. Zhou, X., Benson, K. F., Ashar, H. R., and Chada, K. (1995) Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* **376,** 771–774

90. Dailey, L., and Basilico, C. (2001) Coevolution of HMG domains and homeodomains and the generation of transcriptional regulation by Sox/POU complexes. *J. Cell. Physiol.* **186,** 315–328

91. Remenyi, A., Lins, K., Nissen, L. J., Reinbold, R., Scholer, H. R., and Wilmanns, M. (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17,** 2048–2059

92. Swanson, P. C. (2002) Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. *Mol. Cell. Biol.* **22,** 1340–1351

93. Pierantoni, G. M., Finelli, P., Valtorta, E., Giardino, D., Rodeschini, O., Esposito, F., Losa, M., Fusco, A., and Larizza, L. (2005) High-mobility group A2 gene expression is frequently induced in non-functioning pituitary adenomas (NFPAs), even in the absence of chromosome 12 polysomy. *Endocr. Relat. Cancer* **12,** 867–874

94. Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., Bell, G. W., Otte, A. P., Vidal, M., Gifford, D. K., Young, R. A., and Jaenisch, R. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441,** 349–353

95. Wissmuller, S., Kosian, T., Wolf, M., Finzsch, M., and Wegner, M. (2006) The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors. *Nucleic Acids Res.* **34,** 1735–1744

96. Butteroni, C., De Felici, M., Scholer, H. R., and Pesce, M. (2000) Phage display screening reveals an association between germline-specific transcription factor Oct-4 and multiple cellular proteins. *J. Mol. Biol.* **304,** 529–540

97. Nemeth, M. J., Kirby, M. R., and Bodine, D. M. (2006) Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 13783–13788

98. Nemeth, M. J., Cline, A. P., Anderson, S. M., Garrett-Beal, L. J., and Bodine, D. M. (2005) Hmgb3 deficiency deregulates proliferation and differentiation of common lymphoid and myeloid progenitors. *Blood* **105,** 627–634

99. Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., and Orkin, S. H. (2006) A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444,** 364–368