Lineage-specific interface proteins match up the cell cycle and differentiation in embryo stem cells

Stem Cell Research (2014) 13, 316–328

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/scr

Lineage-specific interface proteins match up the cell cycle and differentiation in embryo stem cells

Angela Re[a,b], Christopher T. Workman[b], Levi Waldron[c], Alessandro Quattrone[a,*], Søren Brunak[b,d,**]

[a] Laboratory of Translational Genomics, Centre for Integrative Biology, University of Trento, Via delle Regole 101, I38123 Trento, Italy
[b] Center for Biological Sequence Analysis, Technical University of Denmark, Kemitorvet, DK2800 Lyngby, Denmark
[c] City University of New York School of Public Health, Hunter College, 2180 3rd Avenue, NY 10035, USA
[d] Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Blegdamsvej 3B, DK2200 Copenhagen, Denmark

Received 11 March 2014; received in revised form 25 July 2014; accepted 26 July 2014
Available online 4 August 2014

Abstract The shortage of molecular information on cell cycle changes along embryonic stem cell (ESC) differentiation prompts an in silico approach, which may provide a novel way to identify candidate genes or mechanisms acting in coordinating the two programs. We analyzed germ layer specific gene expression changes during the cell cycle and ESC differentiation by combining four human cell cycle transcriptome profiles with thirteen in vitro human ESC differentiation studies. To detect cross-talk mechanisms we then integrated the transcriptome data that displayed differential regulation with protein interaction data. A new class of non-transcriptionally regulated genes was identified, encoding proteins which interact systematically with proteins corresponding to genes regulated during the cell cycle or cell differentiation, and which therefore can be seen as interface proteins coordinating the two programs. Functional analysis gathered insights in fate-specific candidates of interface functionalities. The non-transcriptionally regulated interface proteins were found to be highly regulated by post-translational ubiquitylation modification, which may synchronize the transition between cell proliferation and differentiation in ESCs.

Introduction

The process of differentiation of mammalian embryonic stem cells (ESCs) involves an increasing restriction in proliferative capacity (Nichols and Smith, 2009), ending in cell cycle exit in terminally differentiated cells (Coronado et al., 2013; Roccio et al., 2013; Ruiz et al., 2011). For a successful differentiation

* Corresponding author. Fax: +39 0461283096.
** Correspondence to: S. Brunak, Center for Biological Sequence Analysis, Technical University of Denmark, Kemitorvet, DK2800 Lyngby, Denmark. Fax: +45 45931585.
E-mail addresses: alessandro.quattrone@unitn.it (A. Quattrone), brunak@cbs.dtu.dk (S. Brunak).

http://dx.doi.org/10.1016/j.scr.2014.07.008

process, the embryonic transcriptional regulatory programs instructing proliferation should be coordinated with it. However, the regulation of cell cycle related genes during ESC differentiation remains unclear (Roccio et al., 2013). Despite recent progress (Pauklin and Vallier, 2013), whether a change in cell cycle regulation is in itself causative of a change in developmental potential is largely unknown. Instead, much work on cell cycle control in ESCs has focused on features likely associated with the establishment and maintenance of pluripotency.

ESCs have very unusual cell cycle structure, characterized by a short cell division cycle time, truncated G1 and G2 phases, and a large proportion of cells in the S phase (Hindley and Philpott, 2013; Kapinas et al., 2013; Orford and Scadden, 2008; White and Dalton, 2005). Several studies report controversial observations about ESC cycle-specific cyclin-dependent kinase (CDK) activities (Ballabeni et al, 2011; Neganova et al. 2009; Sela et al. 2012). More broadly, several surveillance mechanisms handling genome stability and cell cycle progression are known to operate differently in ESCs (Kapinas et al. 2013; Hussein et al. 2013; Neganova et al., 2011; Becker et al., 2010). The most notable example of this unconventional behavior is the overruling of the restriction (R) point, which is thought to shield ESCs from extrinsic differentiation cues operating during early G1 and to allow ESCs to execute full proliferation (Orford and Scadden, 2008; Sage, 2012; Calder et al., 2013). Supporting this observation, acquisition of the R point control, presumably through the activation of the retinoblastoma-related family of proteins, is an early event in ESC differentiation (Ruiz et al., 2011; Hindley and Philpott, 2013). Moreover, recent advances support the notion that in ESCs the subnuclear reorganization of transcription during cell proliferation is different from that in differentiated cells (Meuleman et al., 2013; Aoto et al., 2006).

Importantly, the unusual cell cycle has also been shown to positively correlate with the pluripotent state, although the molecular mechanisms are not fully understood; for instance, several experiments linked Oct-4, Nanog and Myc to CDKs and their inhibitors (Singh and Dalton, 2009) and to chromosome segregation factors (Nitzsche et al., 2011). Additional evidence that the unconventionally fast cell cycle kinetics in ESCs is associated with their pluripotent state comes from the loss of this behavior upon differentiation (Calder et al., 2013; White and Dalton, 2005) and the reacquisition of it upon reprogramming (Ghule et al., 2011; Egli et al.; 2008).

Despite these advances, an unbiased genome-wide dissection of the relationship between the programs of cell cycle control and ESC differentiation would ideally require the observation of in vitro synchronously differentiating ESCs. Such an experiment is challenging for reasons including heterogeneous mitotic activities across an ESC colony (Jin et al, 2010), the exceedingly rapid ESC cycle and the reported biasing effects of ESC synchronization protocols on cell death and differentiation (Sela et al., 2012; Schneider and d'Adda di Fagagna, 2012; Zhang et al., 2005). Hence, a preliminary in silico approach is an attractive possibility to identify and prioritize genes, pathways and processes for further analysis.

To address this goal, we first assembled a large body of data on the transcriptome dynamics of the human cell cycle and of in vitro ESCs committed to differentiation. In order to identify links between the two programs, we combined transcriptome-level information on periodic genes during the human cell cycle with transcriptome-level information on gene expression in differentiating human ESCs. We then mined a physical interaction network using the corresponding proteins in the two programs to identify a third set of proteins (here named interface proteins) interacting with proteins regulated in expression in both processes. As a result, we produced a genome-wide view of the overlap as well as of direct and indirect molecular interactions between periodic genes during cell cycle and differentially expressed genes in ESC differentiation. We characterized the major classes of proteins operating at the interface between embryo proliferation and differentiation and explore their regulation by post-translational modifications.

Materials and methods

Identification of differentially expressed genes related to ESC differentiation fate

Our transcriptome analysis consisted of four major phases: (a) systematic survey of literature for selection of datasets representative of human ESC differentiation, (b) meta-analysis of differentially expressed genes in relation to the differentiation toward each germ layer, (c) evaluation of the stability for the gene models derived from meta-analysis, and (d) validation in independent datasets.

Dataset assembly

We reviewed the literature and acquired corresponding microarray data (Supplementary Table 1) from the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) or Array Express (http://www.ebi.ac.uk/arrayexpress) databases. Factors for study selection were the following: (1) data availability in human ESC conditions preceding and following the induction of differentiation, (2) minimal number of three replicates for each condition, (3) raw data access, (4) individual sample annotation as to the differentiation protocol and the cellular differentiation fate, and, (5) exclusion of embryoid body samples. Within the selected studies we defined the datasets by ensuring uniformity according to the adopted differentiation protocol and the annotation of differentiated samples by fate. Eleven datasets, probing for transcriptome changes before and after the induction of differentiation, were selected to identify DEX genes, whereas ten independent datasets each sampling multiple time points were set aside for the purpose of validation of the methods for discovering DEX genes. Datasets for DEX gene discovery and for validation were separated by corresponding cellular differentiation fate.

Differential expression meta-analysis

Original data were corrected for background, normalized by the quantile method, summarized to gene level and $\log_2$-transformed within each dataset. We calculated fixed-effects gene regression models within each dataset using the *limma* package (Smyth, 2004) and we synthesized the models across datasets sharing common cellular differentiation fate using the *metafor* R package (Viechtbauer, 2010). Genes were ordered by the estimated coefficients in the

meta-regression models for each cellular differentiate fate. We checked for heterogeneity across the datasets by calculating the percentage of genes where each discovery dataset was found influential by the *metafor* package. The primary analysis was based on fixed-effects models. To additionally address the influence of heterogeneity in the discovery datasets, we compared results from our fixed-effects meta-analysis to those from a random-effects meta-analysis, which models heterogeneity across datasets as random differences between experiments.

Additional analyses

To validate the meta-regression models obtained from each fate-specific meta-analysis, within each validation dataset we calculated fixed-effects gene regression models by using the initial time point as intercept in each multiple contrast corresponding to the additional differentiation time points. P-values associated with each regression model coefficient of the previously identified DEX genes were transformed into False Discovery Rate (FDR) estimates. DEX genes were defined at varying thresholds of the estimated coefficients in the meta-regression models obtained from previous fate-specific meta-analyses. A DEX gene was considered validated if: (a) we identified at least one time point when the gene yields an FDR lower than 0.05 and consistent direction of differential expression in discovery and validation datasets, and (b) if we did not identify any time point when the gene yields an FDR lower than 0.05 and the direction of differential expression was inconsistent with discovery data. We calculated the fraction of DEX genes whose coefficients in the regression models from a validation dataset did not conflict with the coefficients in the meta-regression models previously obtained for each cellular fate. We note that model validation relied on validation datasets annotated to the same fate of discovery datasets. Validation results were summarized by averaging validation performances over all datasets identified for each cellular fate.

To verify output stability to batch effects in each fate-specific meta-analysis, we adjusted the data either for dataset or for platform batch effects by using the *Combat* approach (Johnson et al. 2007) and we carried out fate-specific meta-analysis in each case. We assessed the influence of batch effects by estimating the rate of DEX gene confirmation after correction for batch effects. For stability quantification, DEX genes were defined at varying thresholds of the estimated coefficients in the meta-regression models obtained from the original fate-specific meta-analyses. Batch correction was not helpful, lowering validation rates, and was not used.

Our analysis of validation datasets differed from that of the discovery ones and, in particular, the lists of DEX genes were not derived from validation datasets. To further assess the robustness of the DEX genes identified from the discovery datasets, we used the validation datasets equivalently to the discovery datasets and added a comparative analysis between the DEX genes derived in two ways (Fisher's Exact Test).

Identification of cell cycle periodic genes

We took advantage of a previous, major study (Jensen et al. 2006) that used four microarray expression time courses for an unsupervised, genome-wide screen of genes periodically regulated during the mammalian cell cycle (referred to as PER genes). The 600 periodic cell cycle human genes are available at http://www.cyclebase.org/.

Association between cell cycle periodic genes and ESC differentiation genes

To assess the direct association between differential expression in human ESC differentiation and cell cycle periodicity we tested the significance of the overlap between PER and DEX gene sets using the hypergeometric test for overrepresentation of overlap relative to the overlap expected from two random gene sets selected from all human genes (P-value ≤ 0.01).

We collected protein–protein interactions (PPIs) and PPI-related scores for the proteins corresponding to the PER and DEX genes from the Inweb network (Lage et al. 2007). We selected strictly binary interactions between proteins for which we gathered evidence of expression at the protein level in human ESCs. We derived the human ESC proteome from quantitative mass spectrometry-based proteomics studies (Phanstiel et al., 2011; van Hoof et al., 2009). The relationship between the PER and DEX sets was estimated by calculating the following quantities: the count of direct interactions between the proteins encoded by PER and DEX genes respectively, the count of topological first-order interaction neighbors common to the PER and DEX sets, the count of significant interface (INT) proteins in the Inweb network. Significant interface proteins were defined as proteins at the interface between any pair of PER and DEX proteins which were shown to yield highly overlapping interaction neighbors (Benjamini–Hochberg (BH) adjusted hypergeometric test $P < 1E-03$). For each type of association measure, statistical significance of the estimated measure was empirically assessed by a randomization procedure, based on the generation of 10,000 pairs of randomized PER and DEX sets which approximately preserved the degree distributions of the real PER and DEX sets.

Functional analyses

Differentially regulated genes related to fate-specific ESC differentiation were subjected to functional enrichment analysis of Gene Ontology Biological Process categories in *DAVID* (Huang et al. 2009) by setting False Discovery Rate (FDR) < 0.01. Similar analyses were carried out for the genes resulting to be both PER and DEX genes and for the INT genes.

To draw the functional map of interface versus interfaced functionalities, significant interface proteins were ordered according to the average score of their interactions with the proteins encoded by the PER and DEX gene sets. The ordered proteins were then subjected to Gene Set Enrichment Analysis (GSEA) of Gene Ontology (GO) Biological Process (BP) terms (FDR < 0.25). The next step was to obtain the unshared interaction neighbors for the PER-encoded and DEX-encoded proteins which yielded significant interaction overlaps via proteins annotated to each GSEA-derived GO term. For a GSEA-derived GO term, the identified PER and DEX interaction neighbors were then subjected to Gene Set

Enrichment Analysis (FDR < 0.25) Preliminarily to GSEA, the unshared interaction neighbors of PER and DEX proteins were ordered by the average score of their interactions with the PER and DEX proteins respectively. Leading Edge Analysis was executed in combination with GSEA in order to examine the genes in the leading edge subsets of selected enriched GO categories.

Finally, we estimated the semantic similarity of the resulting GSEA terms to the representative GO terms "cell cycle" and "cell differentiation" by three alternative measures. The ranks obtained by using each measure were compared by the Pearson's correlation coefficient.

Ubiquitylation analysis

PEST regions (Rogers et al., 1986), D-box and KEN-box motifs were predicted into the proteome by using the "epest" and "patmatdb" tools within EMBOSS with default parameters (Rice et al., 2000). We based the identification of putative KEN box and D box instances on the regular expressions K-E-N-X-X-X-N (Pfleger and Kirschner, 2000) and R-X-X-L-X-X-X-X-N (Glotzer et al., 1991), respectively. Protein sequences were retrieved by BioMart; the longest sequence was retained for each protein. Degradation motif instances were mapped onto the proteins corresponding to the PER, DEX and INT genes. The enrichment of the PER, DEX and INT gene sets in putative degradation targets in comparison to the proteome was assessed by the hypergeometric test.

To intersect the PER, DEX and INT gene sets with the data on differential protein ubiquitylation detected in mouse ESC differentiation, we used a quantitative mass spectrometry experiment of global changes in protein ubiquitylation in response to ESC differentiation based both on SILAC and on label-free approaches. We collected proteins differentially ubiquitylated using an FDR of 1% according to either approach. Proteins were classified in more highly detected in the pluripotent or in the differentiated state; 163 proteins were discarded due to conflicting detection of changes by the two approaches. The hypergeometric test was used to assess the enrichment of the PER, DEX and INT gene sets in differentially ubiquitylated proteins, separately for each set in comparison to the proteome. Differentially ubiquitylated proteins were subjected to functional enrichment analysis of GO BP categories by using the hypergeometric test (BH adjusted P < 0.05).

Results

Periodically regulated cell cycle genes overlap with genes differentially regulated during embryonic stem cell differentiation

To systematically probe for cross-talk between transcriptional networks operating in the regulation of ESC differentiation and in the cell cycle, we combined transcriptome data from experiments performed to describe these two programs. The initial compendium was assembled by a systematic review and evaluation of human microarray transcriptome studies profiling human ESC differentiation (Table 1, Fig. 1A). Subsequent extraction of datasets of combined undifferentiated and differentiating samples toward a specific germ layer led to the identification of 21

Table 1 Summary of datasets included in the meta-analysis. Datasets are annotated by the usage in DEX gene discovery or validation, internal identifier, study of origin, ESC differentiation sub-lineage (as per sample annotation at the public microarray database used for retrieval), sampled time points, microarray platform identifier and corresponding PubMed reference (when available).

Usage	Series	Dataset	Differentiation sub-lineage	Time	Platform	PubMed
Discovery	Ect2	GSE34201	Neural stem cell	0,10	GPL6884	22678061
Discovery	Ect3	GSE34201	Neural stem cell	0,10	GPL6884	22678061
Discovery	Ect4	GSE34201	Neural stem cell	0,10	GPL6884	22678061
Discovery	Ect5	GSE34201	Neural stem cell	0,10	GPL6884	22678061
Discovery	Ect8	GSE8590	Neural stem cell	0,21	GPL570	21142452
Discovery	Endl	GSE25557	Definitive endoderm	0,3	GPL6244	21151107
Discovery	End5	GSE16681	Definitive endoderm	0,4	GPL7363	19807270
Discovery	End6	E-MTAB-351	Definitive endoderm	0,3	GPL6883	21245162
Discovery	Mes1	GSE8590	Mesodermal precursor cell	0,21	GPL570	21142452
Discovery	Mes2	GSE8590	Mesodermal precursor cell	0,21	GPL570	21142452
Discovery	Mes3	GSE15257	Renal precursor cell	0,14	GPL6102	20143954
Validation	Ect1	GSE9940	Neural precursor cell	0,10,17	GPL570	n.a.
Validation	Ect10	GSE45223	Melanocyte	0,8,11	GPL10558	n.a.
Validation	Ect6	GSE28633	Mature neural cell	0,9,13,37	GPL6947	21829537
Validation	Ect9	GSE45223	Neural crest cell	0,1,3,6,8,11	GPL10558	n.a.
Validation	End3	GSE25046	Mature hepatocyte cell	0,5,20	GPL6947	21505074
Validation	End8	E-MTAB-467	Definitive endoderm	0,1,2,3	GPL6883	21245162
Validation	End9	E-MTAB-817	Pancreatic cell	0,1,2,5,8,11	GPL6947	n.a.
Validation	Mes4	E-MEXP-3371	Cardiomyocyte	0,7,14	GPL6884	22020065
Validation	Mes5	E-MTAB-1510	Endothelial cell	0,2,4,10	GPL6947	23618383
Validation	Mes6	E-MTAB-781-464	Smooth muscle cell	0,5,17	A-MEXP-2072	22252507

A

FATE	DISCOVERY (two time-points)			VALIDATION (multiple time-points)		
	STUDY	DATA SET	SAMPLE	STUDY	DATA SET	SAMPLE
ECTODERM	2	5	36	3	4	48
ENDODERM	3	3	20	3	3	68
MESODERM	2	3	18	3	3	44

B

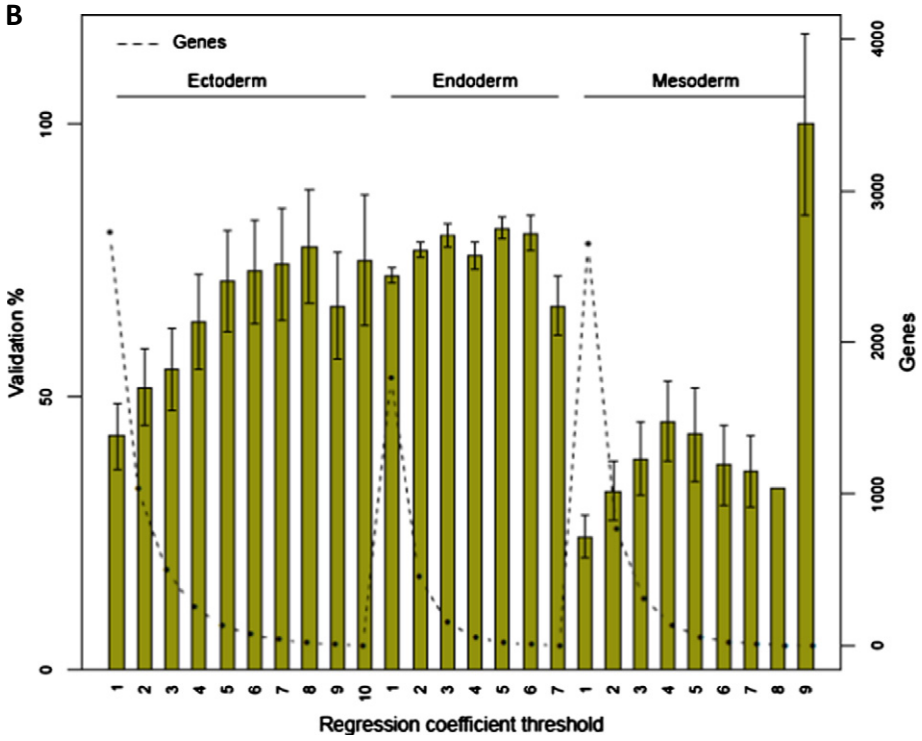

Figure 1 Summary of meta-analysis fixed-effects regression models of differential expression in human ESC differentiation. (A) Summary of samples and datasets derived from studies selected for meta-analysis of differentially expressed genes in human ESC differentiation toward distinct cellular fates. Datasets consisting of two time points, which were chosen to identify differentially expressed genes, were separated from datasets consisting of multiple time points, which served for independent validation purpose. (B) Validation results of fixed-effects regression models synthesized across datasets for each cellular differentiation fate. Validation was conducted by assessment of model consistency in independent datasets consisting of multiple time points. Error bars represent the standard errors of the average validation rates across fate-specific validation datasets. Validation analysis was conducted at varying threshold on meta-analytic gene regression coefficients by cellular differentiation fate. The dotted line shows the number of genes retained at each threshold.

Table 2 Summary of fixed-effects differential expression and validation analysis. From left to right: number of DEX genes from lineage-wise meta-analyses of discovery datasets; number of genes where a dataset contributing to a lineage-wise meta-analysis was found influential in model fit; number of DEX genes validated in at least one up to the total number of validation datasets available per lineage.

Lineage	Fixed effects DEX	% genes where i dataset is influential					No. genes validated in N datasets			
Ect	2722	Ect2	Ect3	Ect4	Ect5	Ect8	N ≥ 1	N ≥ 2	N ≥ 3	N ≥ 4
		2.17	2.17	2.03	3.81	10.31	2032	1226	533	0
End	1768	Endl	End5	End6	n.a.	n.a.	N ≥ 1	N ≥ 2	N ≥ 3	
		4.47	4.65	7.47			1666	1302	725	
Mes	2642	Mesl	Mes2	Mes3	n.a.	n.a.	N ≥ 1	N ≥ 2	N ≥ 3	
		4.48	4.65	7.47			1372	290	4	

datasets. Eleven datasets consisted of two time points which gauged transcriptome changes before/after induction of differentiation; these *discovery datasets* were used to identify differentially expressed (DEX) genes. Ten datasets consisted of multiple time points; these independent *validation datasets* were set aside to validate the methods for identifying DEX genes. We emphasize that we did not use these datasets as a filter for the discovered DEX genes. Fate-specific differential expression results were combined across discovery datasets by fixed-effects meta-analysis. For the purpose of validation, we calculated, in each independent validation dataset, the fraction of DEX genes whose differential expression was observed at an FDR lower than 0.05 and whose direction of differential expression did not conflict with the combined coefficient of linear regression from discovery datasets (Fig. 1B). The validation percentages were satisfactory for each differentiation fate, albeit lower for the mesodermal fate (~50% at the middle threshold instead of 75%). Furthermore, inspection of DEX gene validation by differentiation stage revealed high percentages of validated DEX genes at the time points in common to discovery/validation datasets (Supplementary Fig. 1). The complete analysis process is summarized in Supplementary Fig. 2.

For the purpose of identifying genes regulated both by cell cycle and during in vitro ESC differentiation, genes with a differential expression (combined over all discovery datasets by meta-analysis) above a fold-change of two and significance level of 0.01 were considered differentially expressed (DEX) genes. Ectodermal, endodermal and mesodermal DEX genes amounted to 2722, 1768 and 2642 genes respectively, with pairwise Jaccard indices ranging from 0.17 to 0.23 (Fisher's Exact Test, P-value < E−16) and 406 genes common to all fates (Table 2, Supplementary Table 2). Although some evidence of heterogeneity across discovery datasets was detected, we confirmed that it did not affect the stability of the DEX gene identification (Supplementary

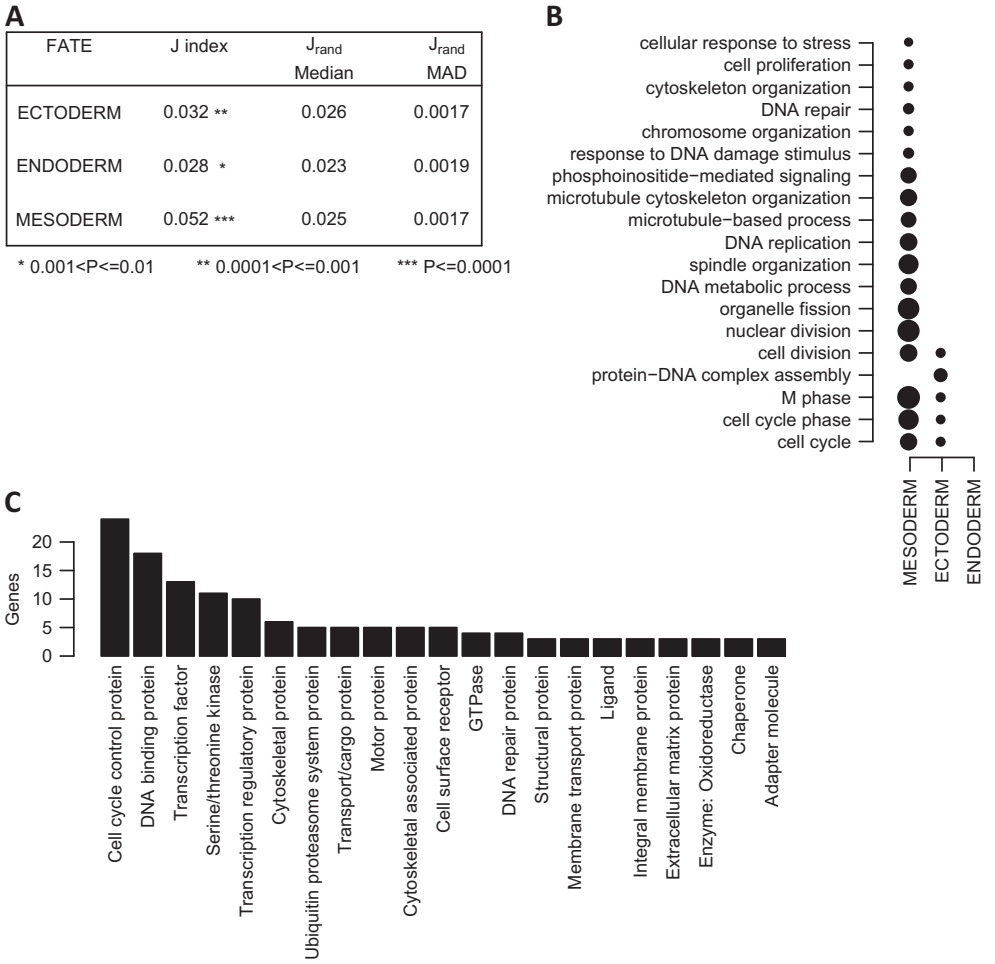

Figure 2 Periodically regulated cell cycle genes overlap with genes differentially regulated during embryonic stem cell differentiation. (A) Table showing the Jaccard index between PER and DEX genes identified per cellular differentiation fate. The accompanying '*' symbols represent the statistical significance of the overlap according to the legend. The table also reports the median Jaccard indices and median absolute deviations which were obtained by 10,000 randomizations of PER and DEX gene sets. (B) Enrichment of overlapping PER and DEX genes in Gene Ontology (GO) Biological Processes (BP) categories by cellular differentiation fate (FDR < 0.01). Dot size reflects the Jaccard index between overlapping genes and GO BP gene sets. (C) Breakdown of overlapping PER and DEX genes by molecular class.

information, Supplementary Tables 3, 4). Functional enrichment analysis of differentially regulated genes revealed phenotypic changes consistent with the cellular differentiation fates, supporting the high quality of DEX genes yielded by our analysis (Supplementary Fig. 3).

We then derived a quality-controlled set of 600 periodically regulated (PER) genes in the human cell cycle from a curated, comprehensive repository of cell cycle experiments and analysis results (Gauthier et al, 2010). The integration between cell cycle progression and ESC differentiation could potentially be manifested by genes co-regulated by the two programs, i.e., shifting from oscillation during cell proliferation toward a polarized increase or decrease during differentiation. To verify this hypothesis, we assessed the overlap between cell cycle regulated genes and genes regulated during in vitro ESC differentiation, and found that this overlap was greater than expected by chance for each differentiation fate (hypergeometric test P-value ≤ 0.01, Fig. 2A). Functional enrichment analysis of the overlapping genes primarily highlighted cell cycle related functionalities, such as DNA replication and cell division (Fig. 2B). The overlapping genes were mainly found to be cell cycle control genes including cyclins (CCNE2, CCNB1, CCNA2), replication factors (PCNA, CDC6, MCM6, CKS2, CKS1B) and mitotic spindle assembly factors such as TACC3 and CDC20. The overall range of functions was broader. Indeed the genes detected in each fate-specific overlap between PER and DEX sets included transcription factors, GTPases, heat shock proteins, and cytoskeletal and transport proteins (Fig. 2C). Remarkably, dual roles in cell proliferation and differentiation were previously described for several overlapping genes identified in our analysis (Supplementary Table 5) such as the DNA replication inhibitor GMNN, which is known to regulate embryonic transition from pluripotency to early multi-lineage commitment (Lim et al., 2011), and the H1 histone family member H1F0, whose depletion was shown to affect mitotic chromosome architecture and segregation (Maresca et al., 2005) and to reduce developmental gene expression (Zhang et al., 2012). Results were robust to the use of alternative regression models (Supplementary information, Supplementary Figs. 4, 5) and choice of datasets for DEX gene identification (Supplementary information, Supplementary Table 6, Supplementary Figs. 6, 7). Furthermore, as a major result of our careful identification of fate-specific DEX genes, we noted that the significance of the overlap between cell cycle regulated genes and genes regulated during in vitro ESC differentiation depended on the differentiation fate (Figs. 2A, B). To clarify the overlap fate specificity, we conducted fixed-effects meta-regression analysis using samples from all lineages (Supplementary Fig. 8). This approach showed lack of significant overlap and lack of functional enrichment for the overlapping genes and did not support the possibility that gene co-regulation during the processes of cell cycling and of ESC differentiation may be fate-unspecific. Rationalizing the highly significant overlap observed between the cell cycle regulated genes and mesodermal DEX genes deserves future investigation. Our results indicate that gene co-regulation during the processes of cell cycling and of ESC differentiation may be fate-specific.

We next extended our analysis to explore the potential network infrastructure which may help to integrate the programs of cell proliferation and differentiation by means of genes which are neither periodical nor differentially expressed during ESC commitment. To verify this possibility we used Inweb, an updated high quality human protein interaction map (Lage et al, 2007) consisting of protein–protein interactions, and we extracted interaction neighbors of the proteins coded for by the two PER and DEX gene sets. Some of these bridging interface genes might be used by ESCs to couple expression changes determining cell cycle dynamics with those promoting exit from pluripotency and induction of differentiation.

Interface genes linking cell cycle and embryonic stem cell differentiation code for central network proteins

From Inweb we selected direct physical protein–protein interactions (PPIs) where both proteins were required to be expressed in human ESCs; PPIs finally involved 171, 163 and 175 proteins encoded by PER genes and 593, 359 and 852 proteins encoded by DEX genes in the ectodermal, endodermal and mesodermal fates, respectively. By approximate node degree-preserving randomization of the PER and DEX sets, network analysis revealed a total number of direct interactions ranging from 407 to 1483 according to the cellular differentiation fate. Direct interaction counts did not indicate a strong tendency for these two sets to directly interact with each other more than or less than expected by chance (Supplementary Fig. 9).

Therefore, we sought to better understand the potential cross-talk between the PER and DEX gene sets by analyzing their interface genes, i.e. the direct interaction neighbors common to the proteins encoded by the PER and DEX genes. In doing so, the PER and DEX sets were found to share a total of INT genes ranging from 1229 to 1458 according to the cellular differentiation fate. Interestingly, by approximate node degree-preserving randomization of the PER and DEX sets, the number of INT genes was found to be significantly fewer than would be expected by chance (underrepresentation P-values ranged from 1E−04 to 2E−03). Moreover, the network centrality of the INT genes was found to be significantly higher in comparison to the centrality measured from similarly defined interface genes between randomized PER and DEX sets (interface degree P-value = 1E−04 irrespective of the cellular differentiation fate). Both the unexpectedly low number of INT genes and their tendency to code for central node proteins in the protein interaction network supported their interface status.

In line with this observation, we focused on the PER and DEX gene pairs featuring significantly highly overlapping interaction neighbors (BH adjusted hypergeometric test P-value $<$ 1E−03) which we refer to as to the interface (INT) genes (Supplementary Fig. 9). The INT genes derived from fate-specific analysis amounted to 1290 out of which 513 INT genes were identified in all fates whereas 184, 74 and 139 INT genes were uniquely identified in the ectodermal, endodermal and mesodermal cellular fates, respectively.

Unanticipated gene functionalities are overrepresented at the interface between cell cycle and embryonic stem cell differentiation

Since previous fate-specific analyses identified both fate-unspecific and fate-specific INT proteins, we separately

conducted functional enrichment analyses (FDR < 0.01) of the fate-unspecific INT proteins and of the INT proteins uniquely identified for each fate. Overrepresented functionalities in the common INT proteins included DNA and RNA metabolism related processes, nucleocytoplasmic transport, protein complex assembly, cytoskeleton organization and cell cycle (Table 3). Functional analyses of fate-specific INT proteins provided fairly distinct results, thus extending the fate-specificity observed in the overlapping PER and DEX genes to the INT proteins. The ectoderm-specific INT proteins confirmed the enrichment in functionalities related to RNA metabolism; in particular, the ectoderm-specific INT proteins were found to regulate RNA splicing (e.g. RBM28, PTBP2, several heterogeneous nuclear ribonucleoproteins) and RNA surveillance (EXOSC2, LSM3, UPF3B). The mesoderm-specific INT proteins instead were enriched in chromatin modeling proteins among which L3MBTL2 whose interaction with a Polycomb Repressive Complex 1 (PCR1)-related complex is tightly linked to its essential requirement in the control of ESC proliferation and differentiation programs (Qin et al., 2012) and KDM5B which regulates H3K4 methylation at developmental genes during ESC differentiation (Kidder et al., 2014). Finally, the endoderm-specific INT proteins were not enriched in any fate-specific functionality (Supplementary Table 7). Our results therefore suggest insights into the overall structure of fate-specificity in the coordination between cell cycle control and ESC differentiation.

Furthermore, we adopted an alternative functional prioritization of INT proteins which explicitly accounted for the scores of the interactions by which the INT proteins were found to connect PER and DEX proteins to each other. More precisely, we applied gene set enrichment analysis (GSEA) (Subramanian et al., 2005) to the INT proteins ordered by the average score of their interactions involving proteins encoded by PER and DEX genes (FDR < 0.25).

In the following, we selected one of the most interesting INT functional categories to suggest a systematic means of using interaction data to study the PER and DEX related functionalities which may be connected by similar functionally annotated INT proteins. For an INT functional annotation, we selected the PER and DEX proteins connected by the annotated INT proteins and applied GSEA to the interaction partners that were not found to be in common between the PER and DEX proteins and that were ordered by the scores of the interactions with PER and DEX proteins, respectively. The functional analysis process is summarized in Fig. 3.

Functionalities overrepresented in INT proteins were prioritized in the order of relatedness to cell cycle and cell differentiation (Supplementary Table 8) by applying several semantic similarity scores with highly consistent results (Supplementary Table 9). An unexpected interface functional category was nuclear transport, which was due to the presence of 26 INT proteins at the interface between 44 PER and 153 DEX proteins. These interface proteins included nucleoporins (NUPs), transport receptors of the importin α and β families and RNA binding proteins, such as HNRNPA1, SAM68 and RAE1 which facilitate transport through the nuclear membrane (Supplementary Table 10). Interestingly, further examination of the identified NUPs, which included both members of the NUP107-160 nuclear pore sub-complex (e.g. NUP133, NUP160) and nuclear ring components such as NUP214, uncovered an emerging dual role in cell cycle (e.g. spindle assembly) (Mishra et al., 2010) and in development related events (Lupu et al., 2008; Sapkota et al., 2007). Additional characterization of INT proteins in relation to dual functional roles in the regulation of cell cycle and cell differentiation is summarized in Supplementary Table 10.

Leading edge analysis of GO categories overrepresented in the 254 PER neighbors isolated 36 genes the majority of which were protein kinases, phosphatases and ubiquitin-conjugating proteins, with known roles in mitotic cell cycle control (PKN2, UBE2C, UBE2D3 and UBE4B) and in signaling pathways such as the RAF–MEK–ERK (Von Kriegsheim et al.,

Table 3 Lineage-specificity of interface protein functionalities. The table reports the FDR values of the GO Biological Process categories which resulted overrepresented in the interface proteins common to all cellular fates or in the fate-specific interface proteins.

GO BP term	Ectoderm interface specific	Endoderm interface specific	Mesoderm interface specific	Common interface
Chromosome organization	n.a.	n.a.	n.a.	1.06E−15
mRNA processing	4.46E−05	n.a.	n.a.	1.28E−15
Cell cycle	n.a.	n.a.	n.a.	2.35E−14
Translation	n.a.	n.a.	n.a.	1.89E−13
Macromolecular complex assembly	n.a.	n.a.	n.a.	7.66E−13
Cytoskeleton organization	n.a.	n.a.	n.a.	3.95E−10
Protein complex assembly	n.a.	n.a.	n.a.	8.87E−08
DNA metabolic process	n.a.	n.a.	n.a.	3.06E−06
Ribosome biogenesis	n.a.	n.a.	n.a.	0.0003
Cell division	n.a.	n.a.	n.a.	0.0007
Protein targeting	n.a.	n.a.	n.a.	0.0007
Nucleocytoplasmic transport	n.a.	n.a.	n.a.	0.0031
Protein folding	n.a.	n.a.	n.a.	0.0055
RNA splicing	6.61E−06	n.a.	n.a.	0.0000
Chromatin modification	n.a.	n.a.	0.0064	n.a.

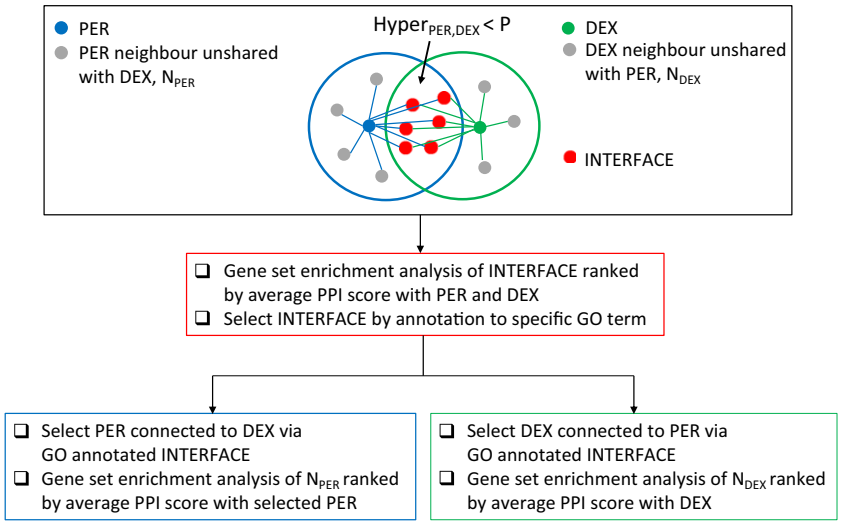

Figure 3 Functional analysis flowchart. Our analysis combines the following major phases: (a) ordering of significant INT proteins by the average score of the interactions with PER- and DEX-encoded proteins, (b) identification of INT functional categories by gene set enrichment analysis of ordered INT proteins, (c) identification of INT proteins annotated to a functional category and identification of the PER, DEX protein pairs with significant interaction neighbors overlap via these INT proteins, and (d) gene set enrichment analyses of the interaction neighbors of PER proteins that are not shared with DEX proteins and vice versa. Unshared interaction neighbors are ordered by the average score of their interactions with PER or DEX proteins.

2006) pathway, the Janus-activated kinase (JAK) and SRC family kinase (SFK) pathways (Shields et al., 2008) (Supplementary Table 11).

A similar analysis of overrepresented GO categories in the 1255 DEX neighbors identified 153 genes covering a broad range of functions and including, in particular, cell cycle genes and developmental genes (Supplementary Table 12). Interestingly, the cell cycle genes were previously shown to be involved in development either by regulating transcription factors, which are responsible for maintaining ESC identity (Card et al. 2008; Deshpande et al. 2009) or promoting differentiation (Morawski et al. 2013), or by regulating signaling pathways, such as the WNT pathway, which controls cell fate determination (Pera and Tam, 2010). Developmental genes predominantly consisted of transcription factors (e.g. FOXO4, GLI2, RELA, SALL1, SOX15 and STAT3) and co-factors (e.g. FHL3, LDB1, SMARCA1), and were accompanied by a number of protein kinases, membrane-associated proteins and small GTPases. Our results, therefore, suggest the intriguing possibility that regulated nuclear-cytoplasmic transport affords mechanic and/or regulatory activities (Raices and D'Angelo, 2012) within pathways triggered by external signals in order to reconcile the demands for embryonic stem cell division and differentiation.

Post-translational modification is a frequent regulatory feature of interface proteins

By the way in which they were identified, we know that the INT genes are not transcriptionally regulated during either the cell cycle or ESC differentiation. We therefore analyzed the INT proteins in terms of their post-translational modifications, which could explain how they regulate the assembly or disassembly of complexes between the proteins encoded by PER and DEX genes. Protein ubiquitylation can modify protein stability, protein subcellular localization (Berlin et al., 2010) or the affinity of protein–protein interactions (Sundd, 2012; Markin et al., 2010). Furthermore, ubiquitylation is known to influence a variety of cell cycle aspects including the cell cycle machinery and cell cycle checkpoints. Therefore, we set out to investigate ubiquitin-dependent regulation in the INT proteins.

PEST sequences, D boxes and KEN boxes, which are the most prominent motifs recognized by ubiquitin ligase complexes, were predicted in the proteome and subsequently mapped onto the protein sequences corresponding to PER, DEX and INT genes. By independently contrasting each set against the proteome, we confirmed the known regulatory role of ubiquitylation in protein coded by PER genes (Jensen et al., 2006) (Fig. 4A). Nonetheless, a more important finding was the extension of substantial enrichment in ubiquitylation motifs to the INT proteins relatively to whole proteome, irrespectively of the cellular fate context where INT proteins were defined (Fig. 4A). By closer inspection, the KEN box motif was found to be uniquely overrepresented in the PER set, the D-box motif in the DEX set whereas the INT set was enriched in PEST regions.

To further confirm the results of the computational motif analysis in the context of ESC differentiation, we cross-referenced our results to a quantitative mass spectrometry experiment of global changes in protein ubiquitylation after 4 days of ESC differentiation induced by LIF withdrawal and retinoic acid addition (Buckley et al. 2012). The number of INT proteins with ubiquitylation-dependent regulation evidence resulted to be significantly higher than expected by chance (Fig. 4A). Therefore, both sequence-based prediction and mass spectrometry-based data agree to suggest ubiquitylation as a major transcription-independent regulatory system for the INT proteins.

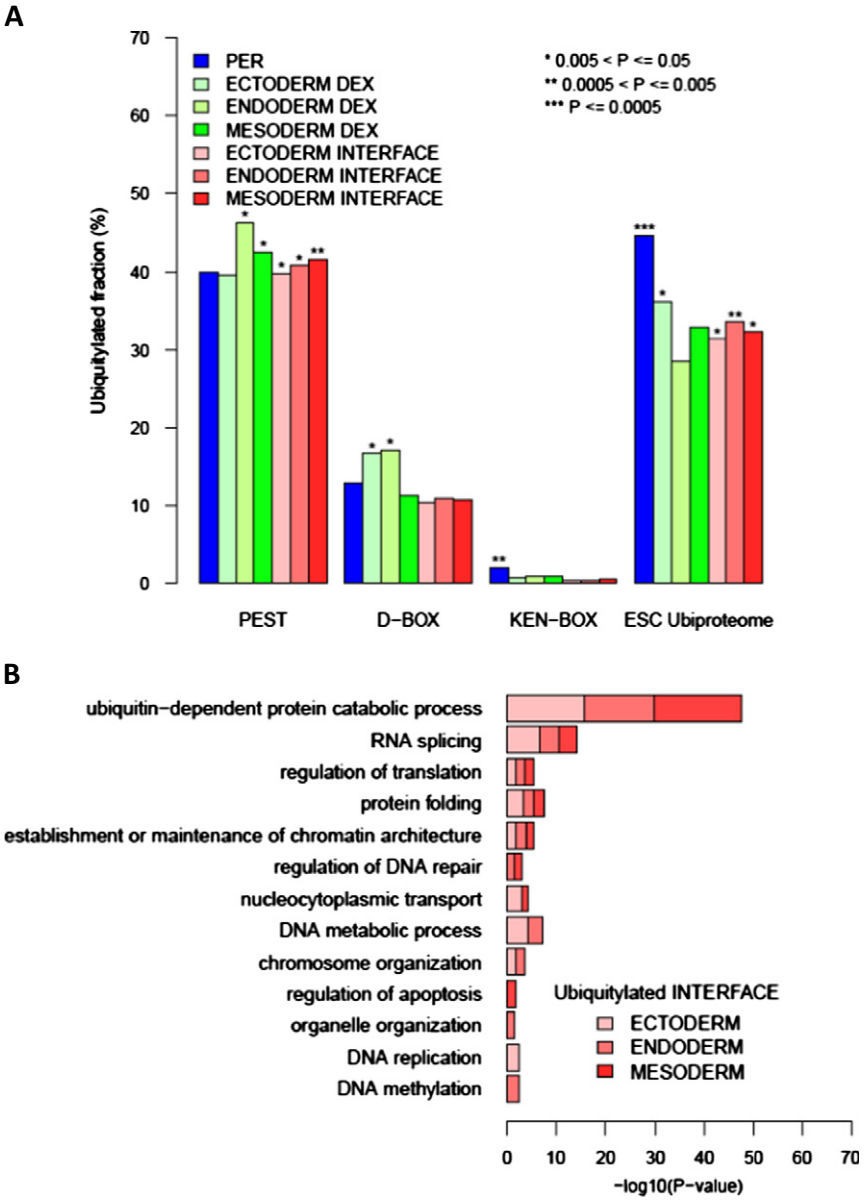

Figure 4 Interface proteins are targeted for ubiquitylation regulation. (A) The bar plot shows the percentages of proteins encoded by the DEG, PER and INT genes and which were found to be ubiquitylated by motif analysis or by experimental evidence in ESC differentiation. The accompanying '*' symbols represent the enrichment statistical significance according to the legend. (B) Functional characterization of interface proteins previously observed to be differentially ubiquitylated. GO categories are ordered by the stability of overrepresentation in lineage-specific analyses.

To study the role of regulated ubiquitylation of the INT proteins in relation to cellular fate, we submitted the INT proteins isolated by the quantitative mass spectrometry experiment to functional enrichment analyses (BH adjusted hypergeometric test P-value < 0.05). The interface functionalities which were invariably identified in relation to each cellular fate occurred to be tightly associated with homeostasis of the proteome (Fig. 4B, Supplementary Table 13). Indeed the identified interface proteins were found to be part of a complex network of cellular machineries which monitor protein life cycle from protein production by translation-dependent regulatory pathways (e.g. nonsense-mediated decay, polysomal recruitment, regulation of translational initiation), to protein folding (e.g. TRiC components), protein localization (in particular nuclear-cytoplasmic transport) up to protein elimination by the ubiquitin proteasome system of which we isolated 11 subunits of the 19S regulator, 10 subunits of the 20S core and 4 E3 ligase complex subunits. Remarkably, we noticed that ubiquitylation-dependent regulation of various interface proteins has been previously shown to have a functional role in regulation of the cell cycle and cell differentiation. This is the case for the deubiquitylating enzyme PSMD14 whose enzymatic activity is essential

for ESC pluripotency (Buckley et al. 2012), the histone demethylase KDM4A, and the transcriptional regulatory TRAPP whose protein turnover associates with replication timing and chromatin condensation during mitotic cell cycle (Van Rechem et al., 2011; Ichim et al. 2014).

3. Discussion

A precise understanding of the relationship between the integrated regulation of ESC proliferation and differentiation might possibly be achieved through a quantitative distinction of the relative contributions of these two programs in the transcriptome variation measured in synchronously differentiating ESCs. However, the achievement of this experimental objective is hampered by several issues, which include the unusually rapid division of these cells (Kapinas et al. 2013), the heterogeneity of cell cycle profiles in self-renewal ESCs and the inefficiency of the synchronization protocols affected by side-effects and by cell death (Jin et al., 2010). Therefore, to shed light on the possibility that transcriptome variations during the cell cycle is systematically coordinated with those promoting ESC transition from self-renewal to differentiation, we devised an in silico approach involving meta-analysis of multiple datasets of in vitro transcriptome data independently measured in synchronously dividing cells or in differentiating ESCs. First we attained a high-quality resource of fate-specific differentially regulated genes in in vitro human ESC differentiation by taking advantage of synthesized results from meta-analysis of previously unconnected datasets. The cell cycle periodically regulated genes were acquired from previous analysis of human cell cycle transcriptome data in somatic cell culture models. Such data represent certainly an approximation of the cell cycle in ESCs. Since a systematic assay of cell cycle transcriptome changes in ESCs is not available, the PER genes, albeit clearly imperfect, are used as a proxy to derive potential, general relationships between the cell cycle and differentiation. We showed that the genes oscillating during the cell cycle overlap with differentially regulated genes during in vitro ESC differentiation and we showed that the extent of this overlap is differentiation fate-specific. Overlapping genes are robust as they were confirmed in a variety of analyses and represent valuable candidates for future studies, as they can enable the transition from ESC self-renewal to differentiation by shifting from oscillation during cell proliferation toward a polarized increase or decrease during differentiation.

Furthermore, systems-wide examination of network relationships between the PER and DEX sets uncovered a class of INT proteins which share the ability to interact with proteins from the PER and DEX sets and thus help coordinate their activities. The INT proteins tended, in comparison to the proteome in general, to be strongly regulated at the post-translational level, namely by ubiquitylation, rather than at the transcriptional level. Further studies may explore the possibility of other PTMs to be significantly overrepresented across the INT proteins.

The ontological functional analysis of the INT proteins supported fate specificity in the INT mechanisms integrating the cell cycle and ESC differentiation. Furthermore, cellular processes which were not obviously related to either the cell cycle or to cell differentiation were highlighted. A notable example of unexpected INT proteins was provided by the identification of several proteins of the nuclear-cytoplasmic transport. Selective nuclear transport of RNAs and regulatory factors is the most obvious route to help coordinate gene activity shifts underlying cell cycle control or differentiation. Importantly, context-dependent transport paths through the nuclear pore complex require distinct transport factors supporting potential regulatory roles for these proteins. For example, a previous study indicated that switching the subtype of importin-α, an INT protein, regulates mouse ESC differentiation through the selective nuclear import of the transcription factors Oct3/4, BRN2 and SOX2 (Yasuhara et al. 2007).

Furthermore, we observed INT proteins clearly connected to cell cycle and differentiation. For instance, cyclin D1 (CCND1), which restricts the activity of Smad2/3 resulting in a switch from endoderm to neuroectoderm potential (Pauklin and Vallier, 2013), was an ectoderm and mesoderm but not endoderm INT protein. CDK4 and CDK6, which were activated by CCND1, were correctly identified among CCND1 protein interaction neighbors and, interestingly, CDK6 was up-regulated in ectoderm-fated ESCs.

This systematic work can be viewed as a source of hypotheses which may drive future experiments in stem cell biology. It is obvious, however, that attempts to control and manage stem cell differentiation also may benefit from it. Indeed, understanding how cell cycle mediates the transition between the transcriptional programs generating differentiated cells will eventually allow us to progress in the controlled ex vivo production of therapeutically relevant cell types. Along this line some recent experiments have shown that both directed differentiation of ESCs and reprogramming of somatic cells can be enhanced by reliable methods for enriching ESCs at specific cell cycle phases (Chetty et al, 2013). Recently, cell cycle manipulations of ESCs, albeit in mouse, were found to strongly affect differentiation (Li and Kirschner, 2014). It is therefore possible that the insights provided by our analysis could lead to practical applications in regenerative medicine.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scr.2014.07.008.

Acknowledgments

This work received funding from the Autonomous Province of Trento (40101746), the European Union's Seventh Framework Programme for Research, technological development and demonstration under grant agreement no PCOFUND-GA-2008-336070 and the Novo Nordisk Foundation.

References

Aoto, T., Saitoh, N., Ichimura, T., Niwa, H., Nakao, M., 2006. Nuclear and chromatin reorganization in the MHC-Oct3/4 locus at developmental phases of embryonic stem cell differentiation. Dev. Biol. 298, 354–367.

Ballabeni, A., Park, I.H., Zhao, R., Wang, W., Lerou, P.H., Daley, G. Q., Kirschner, M.W., 2011. Cell cycle adaptations of embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 108, 19252–19257.

Becker, K.A., Ghule, P.N., Lian, J.B., Stein, J.L., van Wijnen, A.J., Stein, J.S., 2010. Cyclin D2 and the CDK substrate p220(NPAT) are required for self-renewal of human embryonic stem cells. J. Cell. Physiol. 222, 456–464.

Berlin, I., Schwartz, H., Nash, P.D., 2010. Regulation of epidermal growth factor receptor ubiquitylation and trafficking by the USP8·STAM complex. J. Biol. Chem. 285, 34909–34921.

Buckley, S.M., Aranda-Orgilles, B., Strikoudis, A., Apostolou, E., Loizou, E., Moran-Crusio, K., Farnsworth, C.L., Koller, A.A., Dasgupta, R., Silva, J.C., et al., 2012. Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. Cell Stem Cell 11, 783–798.

Calder, A., Roth-Albin, I., Bhatia, S., Pilquil, C., Lee, J.H., Bhatia, M., Levadoux-Martin, M., McNicol, J., Russell, J., Collins, T., 2013. Lengthened G1 phase indicates differentiation status in human embryonic stem cells. Stem Cells Dev. 22, 279–295.

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., Archer, T.K., 2008. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol. Cell. Biol. 28, 6426–6438.

Chetty, S., Pagliuca, F.W., Honore, C., Kweudjeu, A., Rezania, A., Melton, D.A., 2013. A simple tool to improve pluripotent stem cell differentiation. Nat. Methods 10, 553–556.

Coronado, D., Godet, M., Bourillot, P.Y., Tapponnier, Y., Bernat, A., et al., 2013. A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. Stem Cell Res. 10, 118–131.

Deshpande, A.M., Dai, Y.S., Kim, Y., Kim, J., Kimlin, L., Gao, K., Wong, D.T., 2009. Cdk2ap1 is required for epigenetic silencing of Oct4 during murine embryonic stem cell differentiation. J. Biol. Chem. 284, 6043–6047.

Egli, D., Birkhoff, G., Eggan, K., 2008. Mediators of reprogramming: transcription factors and transitions through mitosis. Nat. Rev. Mol. Cell Biol. 9, 505–516.

Gauthier, N.P., Jensen, L.J., Wernersson, R., Brunak, S., Jensen, T. S., 2010. Cyclebase.org: version 2.0, an updated comprehensive, multi-species repository of cell cycle experiments and derived analysis results. Nucleic Acids Res. 38, D699–D702.

Ghule, P.N., Medina, R., Lengner, C.J., Mandeville, M., Qiao, M., Dominski, Z., Lian, J.B., Stein, J.L., van Wijnen, A.J., Stein, G.S., 2011. Reprogramming the pluripotent cell cycle: restoration of an abbreviated G1 phase in human induced pluripotent stem (iPS) cells. J. Cell. Physiol. 226, 1149–1156.

Glotzer, M., Murray, A.W., Kirschner, M.W., 1991. Cyclin is degraded by the ubiquitin pathway. Nature 349, 132–138.

Hindley, C., Philpott, A., 2013. The cell cycle and pluripotency. Biochem. J. 451, 135–143.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Hussein, S.M., Elbaz, J., Nagy, A.A., 2013. Genome damage in induced pluripotent stem cells: assessing the mechanisms and their consequences. Bioessays 35, 152–162.

Ichim, G., Mola, M., Finkbeiner, M.G., Cros, M.P., Herceg, Z., Hernandez-Vargas, H., 2014. The histone acetyltransferase component TRRAP is targeted for destruction during the cell cycle. Oncogene 33, 181–192.

Jensen, L.J., Jensen, T.S., de Lichtenberg, U., Brunak, S., Bork, P., 2006. Co-evolution of transcriptional and post-translational cell-cycle regulation. Nature 443, 594–597.

Jin, Q., Duggan, R., Dasa, S.S., Li, F., Chen, L., 2010. Random mitotic activities across human embryonic stem cell colonies. Stem Cells Dev. 19, 1241–1248.

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127.

Kapinas, K., Grandy, R., Ghule, P., Medina, R., Becker, K., Pardee, A. , Zaidi, S.K., Lian, J., Stein, J., van Wijnen, A., et al., 2013. The abbreviated pluripotent cell cycle. J. Cell. Physiol. 228, 9–20.

Kidder, B.L., Hu, G., Zhao, K., 2014. KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. Genome Biol. 15, R32.

Lage, K., Karlberg, E.O., Størling, Z.M., Olason, P.I., Pedersen, A.G., Rigina, O., Hinsby, A.M., Tümer, Z., Pociot, F., Tommerup, N., et al., 2007. A human phenome–interactome network of protein complexes implicated in genetic disorders. Nat. Biotechnol. 25, 309–316.

Li, V., Kirschner, M.W., 2014. Molecular ties between the cell cycle and differentiation in embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 111, 9503–9508.

Lim, J.W., Hummert, P., Mills, J.C., Kroll, K.L., 2011. Geminin cooperates with Polycomb to restrain multi-lineage commitment in the early embryo. Development 138, 33–44.

Lupu, F.I., Alves, A., Anderson, K., Doye, V., Lacy, E., 2008. Nuclear pore composition regulates neural stem/progenitor cell differentiation in the mouse embryo. Dev. Cell 14, 831–842.

Maresca, T.J., Freedman, B.S., Heald, R., 2005. Histone H1 is essential for mitotic chromosome architecture and segregation in *Xenopus laevis* egg extracts. J. Cell Biol. 169, 859–869.

Markin, C.J., Saltibus, L.F., Kean, M.J., McKay, R.T., Xiao, W., Spyracopoulos, L., 2010. Catalytic proficiency of ubiquitin conjugation enzymes: balancing pK(a) suppression, entropy, and electrostatics. J. Am. Chem. Soc. 132, 17775–17786.

Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., van Steensel, B., 2013. Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res. 23, 270–280.

Mishra, R.K., Chakraborty, P., Arnaoutov, A., Fontoura, B.M., Dasso, M., 2010. The Nup107-160 complex and gamma-TuRC regulate microtubule polymerization at kinetochores. Nat. Cell Biol. 12, 164–169.

Morawski, P.A., Mehra, P., Chen, C., Bhatti, T., Wells, A.D., 2013. Foxp3 protein stability is regulated by cyclin-dependent kinase 2. J. Biol. Chem. 288, 24494–24502.

Neganova, I., Zhang, X., Atkinson, S., Lako, M., 2009. Expression and functional analysis of G1 to S regulatory components reveals an important role for CDK2 in cell cycle regulation in human embryonic stem cells. Oncogene 28, 20–30.

Neganova, I., Vilella, F., Atkinson, S.P., Lloret, M., Passos, J.F., von Zglinicki, T., O'Connor, J.E., Burks, D., Jones, R., Armstrong, L., et al., 2011. An important role for CDK2 in G1 to S checkpoint activation and DNA damage response in human embryonic stem cells. Stem Cells 29, 651–659.

Nichols, J., Smith, A., 2009. Naive and primed pluripotent states. Cell Stem Cell 4, 487–492.

Nitzsche, A., Paszkowski-Rogacz, M., Matarese, F., Janssen-Megens, E.M., Hubner, N.C., Schulz, H., de Vries, I., Ding, L., Huebner, N., Mann, M., et al., 2011. RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. PLoS One 6, e19470.

Orford, K.W., Scadden, D.T., 2008. Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. Nat. Rev. Genet. 9, 115–128.

Pauklin, S., Vallier, L., 2013. The cell-cycle state of stem cells determines cell fate propensity. Cell 155, 135–147.

Pera, M.F., Tam, P.P., 2010. Extrinsic regulation of pluripotent stem cells. Nature 465, 713–720.

Pfleger, C.M., Kirschner, M.W., 2000. The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. Genes Dev. 14, 655–665.

Phanstiel, D.H., Brumbaugh, J., Wenger, C.D., Tian, S., Probasco, M.D., Bailey, D.J., Swaney, D.L., Tervo, M.A., Bolin, J.M., Ruotti, V., et al., 2011. Proteomic and phosphoproteomic comparison of human ES and iPS cells. Nat. Methods 8, 821–827.

Qin, J., Whyte, W.A., Anderssen, E., Apostolou, E., Chen, H.H., Akbarian, S., Bronson, R.T., Hochedlinger, K., Ramaswamy, S., Young, R.A., et al., 2012. The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. Cell Stem Cell 11, 319–332.

Raices, M., D'Angelo, M.A., 2012. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. Nat. Rev. Mol. Cell Biol. 13, 687–699.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16, 276–277.

Roccio, M., Schmitter, D., Knobloch, M., Okawa, Y., Sage, D., et al., 2013. Predicting stem cell fate changes by differential cell cycle progression patterns. Development 140, 459–470.

Rogers, S., Wells, R., Rechsteiner, M., 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science 234, 364–368.

Ruiz, S., Panopoulos, A.D., Herrerias, A., Bissig, K.D., Lutz, M., 2011. A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. Curr. Biol. 21, 45–52.

Sage, J., 2012. The retinoblastoma tumor suppressor and stem cell biology. Genes Dev. 26, 1409–1420.

Sapkota, G., Alarcón, C., Spagnoli, F.M., Brivanlou, A.H., Massagué, J., 2007. Balancing BMP signaling through integrated inputs into the Smad1 linker. Mol. Cell 25, 441–454.

Schneider, L., d'Adda di Fagagna, F., 2012. Neural stem cells exposed to BrdU lose their global DNA methylation and undergo astrocytic differentiation. Nucleic Acids Res. 40, 5332–5342.

Sela, Y., Molotski, N., Golan, S., Itskovitz-Eldor, J., Soen, Y., 2012. Human embryonic stem cells exhibit increased propensity to differentiate during the G1 phase prior to phosphorylation of retinoblastoma protein. Stem Cells 30, 1097–1108.

Shields, B.J., Court, N.W., Hauser, C., Bukczynska, P.E., Tiganis, T., 2008. Cell cycle-dependent regulation of SFK, JAK1 and STAT3 signalling by the protein tyrosine phosphatase TCPTP. Cell Cycle 7, 3405–3416.

Singh, A.M., Dalton, S., 2009. The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. Cell Stem Cell 5, 141–149.

Smyth, G.K., 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3 (1) (Article 3).

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102, 15545–15550.

Sundd, M., 2012. Conformational and dynamic changes at the interface contribute to ligand binding by ubiquitin. Biochemistry 51, 8111–8124.

Van Hoof, D., Muñoz, J., Braam, S.R., Pinkse, M.W., Linding, R., Heck, A.J., Mummery, C.L., Krijgsveld, J., 2009. Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell 5, 214–226.

Van Rechem, C., Black, J.C., Abbas, T., Allen, A., Rinehart, C.A., Yuan, G.C., Dutta, A., Whetstine, J.R., 2011. The SKP1-Cul1-F-box and leucine-rich repeat protein 4 (SCF-FbxL4) ubiquitin ligase regulates lysine demethylase 4A (KDM4A)/Jumonji domain-containing 2A (JMJD2A) protein. J. Biol. Chem. 286, 30462–30470.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36, 1–48.

Von Kriegsheim, A., Pitt, A., Grindlay, G.J., Kolch, W., Dhillon, A.S., 2006. Regulation of the Raf–MEK–ERK pathway by protein phosphatase 5. Nat. Cell Biol. 8, 1011–1016.

White, J., Dalton, S., 2005. Cell cycle control of embryonic stem cells. Stem Cell Rev. 1, 131–138.

Yasuhara, N., Shibazaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., Asally, M., Kamachi, Y., Kondoh, H., Yoneda, Y., 2007. Triggering neural differentiation of ES cells by subtype switching of importin-alpha. Nat. Cell Biol. 9, 72–79.

Zhang, E., Li, X., Zhang, S., Chen, L., Zheng, X., 2005. Cell cycle synchronization of embryonic stem cells: effect of serum deprivation on the differentiation of embryonic bodies in vitro. Biochem. Biophys. Res. Commun. 333, 1171–1177.

Zhang, Y., Liu, Z., Medrzycki, M., Cao, K., Fan, Y., 2012. Reduction of Hox gene expression by histone H1 depletion. PLoS One 7, e38829.