Large Scale Phosphoproteome Profiles Comprehensive Features of Mouse Embryonic Stem Cells*‡

Qing-Run Li‡, Xiao-Bin Xing‡, Tao-Tao Chen¶, Rong-Xia Li‡, Jie Dait, Quan-Hu Sheng‡, Shun-Mei Xin¶, Li-Li Zhul, Ying Jinl, Gang Peil, Jiu-Hong Kang¶**, Yi-Xue Li‡‡, and Rong Zeng‡§§

Embryonic stem cells are pluripotent and capable of unlimited self-renewal. Elucidation of the underlying molecular mechanism may contribute to the advancement of cell-based regenerative medicine. In the present work, we performed a large scale analysis of the phosphoproteome in mouse embryonic stem (mES) cells. Using multiplex strategies, we detected 4581 proteins and 3970 high confidence distinct phosphosites in 1642 phosphoproteins. Notably, 22 prominent phosphorylated stem cell marker proteins with 39 novel phosphosites were identified for the first time by mass spectrometry, including phosphorylation sites in NANOG (Ser-65) and RE1 silencing transcription factor (Ser-950 and Thr-953). Quantitative profiles of NANOG peptides obtained during the differentiation of mES cells revealed that the abundance of phosphopeptides and non-phosphopeptides decreased with different trends. To our knowledge, this study presents the largest global characterization of phosphorylation in mES cells. Compared with a study of ultimately differentiated tissue cells, a bioinformatics analysis of the phosphorylation data set revealed a consistent phosphorylation motif in human and mouse ES cells. Moreover, investigations into phosphorylation conservation suggested that phosphoproteins were more conserved in the undifferentiated ES cell state than in the ultimately differentiated tissue cell state. However, the opposite conclusion was drawn from this conservation comparison with phosphosites. Overall, this work provides an overview of phosphorylation in mES cells and is a valuable resource for the future understanding of basic biology in mES cells. *Molecular & Cellular Proteomics 10: 10.1074/mcp.M110.001750, 1–14, 2011.

From the ‡Key Laboratory of Systems Biology and ¶Laboratory of Molecular Cell Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China and ||Key Laboratory of Stem Cell Biology, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences/Shanghai JiaoTong University School of Medicine, Shanghai 200025, China

Received, June 7, 2010, and in revised form, December 9, 2010
Published, MCP Papers in Press, December 13, 2010, DOI 10.1074/mcp.M110.001750

Want to cite this article? Please look on the last page for the proper citation format.
Phosphoproteome Features of Mouse ES Cells

(hES)\(^1\) cells performed to date identified 10,844 phosphorylation sites, which were detected via the combination of two peptide fragmentation methods (22). In addition, two groups recently identified additional phosphoproteomes and profiled their dynamic activities in hES cells (23, 24). However, several major challenges remain unresolved, such as the low abundance of phosphoproteins in general, dynamic site occupancy, and inherently poor fragmentation of phosphopeptides. In contrast with hES cells, a global view of the mES cell phosphoproteome has not yet been fully achieved. mES cells are the most important models that can be used to investigate the "stemness" characteristics of stem cells; therefore, comprehensive information on the phosphorylation states and specific phosphosites in the mES cell proteome will provide an important resource for stem cell research.

In the current study, we attempted to profile a large phosphoproteome data set from mES cells to further explore their phosphorylation characteristics. Furthermore, we developed a "StemCell Project" web site to retrieve and analyze the data from the study. This information provides an overview of phosphorylation in mES cells and should be a valuable resource for the future understanding of the basic biology in mES cells.

EXPERIMENTAL PROCEDURES

mES Cell Culture—An E14.1 cell strain, a substrain of the E14 strain of ES cells derived from the blastocyst of a 129/Ola strain mouse, was purchased from ATCC. The cells were cultured under feeder-free conditions and maintained in Dulbecco’s modified Eagle’s medium with 15% fetal bovine serum, 0.1 mM β-mercaptoethanol (Sigma), 0.1 mM minimal essential medium nonessential amino acids (Chemicon), and 1000 units/ml leukemia inhibitory factor (Chemicon). The medium was changed every day, and cells were passaged every other day. For expansion, the E14 cells within three passages upon thaw were trypsinized into a single cell suspension and plated into Petri dishes at a density of 1.0 × 10^4 cells/cm². Two days after plating, the cells were harvested for proteomics analysis.

Alkaline Phosphatase Staining and Immunofluorescence—E14.1 cells parallel to harvest were plated into a 24-well dish at a density of 1.0 × 10^4 cells/cm². Two days after plating, cells were stained for alkaline phosphatase, SSEA-1, OCT3/4, SOX2, and NANOG expression as a quality control for pluripotency. Alkaline phosphatase staining was carried out using an Alkaline Phosphatase Detection kit (Chemicon, catalog number SCR004) following the manufacturer’s instructions. For immunofluorescence, cells were washed twice with PBS, fixed with 4% PFA in PBS for 30 min, permeabilized with 0.2% Triton X-100 for 10 min, and blocked with 1% BSA in PBS for 2 h. Cells were then incubated with a primary antibody diluted 1:100 in 1% BSA in PBS for 2 h, washed twice with 1% BSA in PBS, incubated with a FITC- or Texas Red-conjugated secondary antibody diluted 1:100 in 1% BSA in PBS for 2 h, washed twice with 1% BSA in PBS, covered with PBS, and scanned for fluorescence under a microscope. Antibodies used were as follows: SSEA-1 (480, Santa Cruz Biotechnology, Santa Cruz, CA, sc21702), OCT3/4 (C-10, Santa Cruz Biotechnology, sc5279), SOX2 (R&D Systems, MAB2018), and NANOG (R&D Systems, AF2729).

Phosphoprotein Purification—To purify the phosphoproteins from cell lysate, the PhosphoProtein Purification kit from Qiagen (Valencia, CA; catalog number 37101) was applied. The procedures were conducted according to the manufacturer’s instructions. A total of 5 mg of crude protein sample was used for the starting sample, and ~200 μg of phosphorylated proteins was obtained for further analysis.

Strong Anion Exchange (SAX)/TiO\(_2\) Fractionation and Enrichment of Phosphopeptides—SAX chromatography was used for fractionation of phosphopeptides of mES cells. Peptides (2 mg) were dissolved in pH 8.5 buffer and loaded onto an SAX column (1.0 mm × 15 cm, Column Technology Inc.) at a flow rate of 0.3 ml/min. During the loading step, the flow-through fraction was collected. The bound peptides were fractionated by a linear gradient at 0.2 ml/min from pH 8.5 to pH 2.0 in 150 min (7.5 min/fraction), and 20 fractions were collected in total. Based on the UV trace, all the fractions were combined, resulting in five fractions that were used for TiO\(_2\) enrichment. The flow-through was enriched four times using TiO\(_2\).

The phosphopeptide enrichment procedure using TiO\(_2\) beads was described elsewhere (25–27) and was used with some modification. In detail, the TiO\(_2\) beads (GL Sciences, Tokyo, Japan) were preincubated first in 200 μl of loading buffer (65% ACN, 2% TFA, saturated by glutamic acid) for acidification. The fractions of the SAX chromatography were resolved in 200 μl of loading buffer and then incubated with ~2 mg of TiO\(_2\) beads. The peptide-bead slurry was incubated and centrifuged, and the supernatant was discarded. The incubated beads were then washed with 800 μl of wash buffer I (65% ACN, 0.5% TFA) and wash buffer II (65% ACN, 0.1% TFA). The bound peptides were eluted once with 200 μl of elution buffer I (300 mM NH\(_4\)OH, 50% ACN) and twice with 200 μl of elution buffer II (500 mM NH\(_4\)OH, 60% ACN). All the incubation, washing, and elution procedures were performed with end-over-end rotation for 20 min at room temperature. The eluates were dried down and reconstituted in 0.1% FA in H\(_2\)O for MS analysis.

Subcellular Fractionation and Protein Digestion—The stepwise subcellular fractionation procedures were performed using the ProteoExtractTM Subcellular Proteome Extraction kit from Calbi-ochem with little modification. Briefly, adherent feeder-free cultured ES cells were gently washed with 1.0 ml of cold wash buffer twice on ice. Detached cells were carefully transferred to a tube and fractionated as well. With ice-cold extraction buffers I–III, the subcellular fractions (cytosolic and nuclear extractions) were collected for proteomics analysis. Protein concentrations were determined by the method of Bradford using BSA as the standard (Bio-Rad). The tryptic digestion of 400 μg of protein from each fraction was processed as described elsewhere (28).

Phosphatase Treatment—Antarctic phosphatase treatment was performed on total cell lysate according to the manufacturer’s instructions (M0289L, Antarctic phosphatase, New England Biolabs) (29). A total of 200 μg of lysate was redis solved in 20 μl of reaction buffer (supplier-provided), 2 μl of 10% SDS, and 158 μl of H\(_2\)O. After Vortex mixing, the sample was incubated at 37 °C for 2 h with 20 μl of 1 unit/μl Antarctic phosphatase added. Finally, the proteins for Western blot validation were recovered after ultrafiltration at 15,000 × g for 1 h at 4 °C.

Validation by Western Blotting—Equal protein amounts (10 μg) from each of the subcellular fractions were loaded for 12% SDS-PAGE, electrophoresed, transferred to PVDF membranes, and immu-

\(^1\) The abbreviations used are: hES, human embryonic stem; mES, mouse embryonic stem; LIF, leukemia inhibitory factor; SAX, strong anion exchange; MDLC, multidimensional liquid chromatography; RP, reversed-phase; SCX, strong cation exchange; IPI, International Protein Index; FDR, false discovery rate; PTM, post-translational modification; MRM, multiple reaction monitoring; Fmoc, N-(9-fluorenyl) methoxycarbonyl; PMID, PubMed identifier; CDK, cyclin-dependent kinase; JAK, Janus kinase; CAMK, Ca\(^2+\)/calmodulin-dependent protein kinase. REST, RE1 silencing transcription factor.
nobilized with antibodies against Hsp90 and Lamin B (Santa Cruz Biotechnology) followed by horseradish peroxidase-conjugated secondary antibodies. The membranes were subjected to chemiluminescence detection (ECL) according to the manufacturer’s instructions (GE Healthcare). Signals of bands from Western blotting were scanned with PDQuest GS-710 on a flat bed scanner.

Each of 50 µg of total cell lysate with and without treatment by phosphatase was subjected to 8 and 15% SDS-PAGE, and then the proteins in the gel were transferred to a nitrocellulose membrane. The membranes were incubated first with the appropriate primary antibodies (anti-NANOG, M-149, Santa Cruz Biotechnology; anti-REST: 07-579, Upstate, Temecula, CA), respectively, and then incubated with HRP-conjugated secondary antibodies for 45 min. The proteins were detected by enhanced chemiluminescence (ECL Plus, GE Healthcare). Signals of bands from Western blotting were scanned with PDQuest GS-710 on a flat bed scanner.

Yin-Yang Multidimensional Liquid Chromatography (MDLC) On-line Fractionation and Mass Spectrometry Analysis—To achieve an in-depth view of mES cells on global and phosphoproteomics analyses, the strategy for peptide fractionation included three methods on the Yin-yang MDLC system (30). The system involves three subsystems with pH continuous gradient elution: 1) strong cation exchange/reversed-phase (SCX/RP) LC-MS/MS (10 steps), 2) SAX/RP LC-MS/MS (10 steps), and 3) SCX/SAX LC-MS/MS (10 steps each). For the SCX/SAX method, the flow-through fractions of peptide samples were collected after loading onto the SCX column, lyophilized, redissolved, and then fractionated on the SAX/RP LC-MS/MS subsystem. The peptide fractionation was performed on an on-line two-dimensional LC-MS/MS system using the pH continuous on-line gradient developed by our laboratory (31) with some modifications. In detail, two 400–200 µM tryptic peptide mixtures of cytosol/nuclear fractions were dissolved in 80 µl of pH 2.5 buffer (pH 8.5 for SAX) before being loaded on an SCX (or SAX) column (320 µm × 100 mm, Column Technology Inc.) followed by 10 continuous on-line pH gradients from pH 2.5 to pH 8.5 for SCX (pH 8.5 to pH 2.0 for SAX). Each fraction from SCX (or SAX) was eluted to the trap columns (Agilent Technologies) alternately and separated on the analytic column (75 µm × 150 mm, Column Technology Inc.). The reversed-phase gradient for analysis was from 2 to 40% mobile phase (0.1% formic acid in acetonitrile (v/v)) in 160 min at 300 nl/min after the split. Another equivalent peptide mixture of cytosol (or nuclear) fraction was dissolved in 80 µl of pH 2.5 buffer and loaded on the SCX column. The flow-through fraction was collected, lyophilized, and redissolved in 80 µl of pH 8.5 buffer prior to loading on the SAX column. The fractionation on the SCX column and SAX column elution were preformed individually on a two-dimensional LC-MS/MS system as mentioned before. For fractions obtained from subcellular fractionation, the cytosolic and nuclear fractions were analyzed repeatedly by SCX and SAX fractionation. Collectively, there were 40 fractions for the cytosol and nuclear extracts, respectively.

All the fractions from RP separation were analyzed in the LTQ-Orbitrap (ThermoFisher Scientific, San Jose, CA) as described (32). The nanospray ionization source was mounted, and the voltage was set at 1.90 kV. Normalized collision energy was 35%. Data-dependent selection of the 10 most intense ions with collision-induced dissociation from each full scan (performed in the Orbitrap) was selected for MS2 analyses (performed in the ion trap). Dynamic exclusion settings were as follows: repeat count, 2; repeat duration, 30 s; exclusion duration, 90 s. The resolution of the Orbitrap mass analyzer was set at 100,000 (m/z 400) for the precursor ion scans.

Data Collection and Database Search—The peak lists of all 108 raw files used for the database search were produced by MaxQuant software (default parameters, version 1.0.13.13) (33). The generated .par and .msm files were searched using the Mascot search engine (version 2.2.0, Matrix Science) (34) against a concatenated forward and reversed mouse International Protein Index protein sequence database (IP1 mouse, version 3.52) containing 55,303 protein sequences and 262 commonly observed contaminants (111,130 protein entries in total). For the Mascot search, the following parameters were used. Cysteine carboxamidomethylation was selected as a fixed modification, and methionine oxidation, protein N-terminal acetylation, and phosphorylation on serine, threonine, and tyrosine were selected as variable modifications. Up to two missing cleavage points of trypsin were allowed. The precursor ion mass tolerance was 7 ppm, and the fragment ion mass tolerance was 0.5 Da for MS/MS spectra. The database search results (.dat files) were further processed with MaxQuant.

All the identification from contaminants was removed first. For protein identification, two peptides were required in which at least one was unique in the database (35). If the identified peptide sequences of one protein were equal to or contained the peptide set of another protein, then the two proteins were grouped together by MaxQuant and reported as one protein group. The false discovery rates (FDRs) for peptide, protein, and phosphosite (the function “apply site FDR separately” was activated) were all set at 1%; and a minimum length of six amino acids was used for peptides identification. The high confidence phosphosites were filtered based on the PTM scores as described (20). In detail, class I phosphorylation sites are defined by a localization probability of 0.75 and a probability localization score difference greater than or equal to 5 (20). The distributions of the site localization probabilities, Mascot scores, and PTM scores of the identified phosphopeptide are shown in supplemental Fig. 3. All the analyses for phosphorylation were based on class I phosphorylation sites. The identified proteins (supplemental Table 1), (phospho-)peptides (supplemental Tables 2 and 3), and MS/MS spectra for phosphopeptides (supplemental Fig. 5) can also be accessed at our StemCell Project web site. The MS raw data were converted using the PRIDE Converter tool (36) and can be retrieved from the Proteomics Identifications (PRIDE) database (http://www.ebi.ac.uk/pride) (37) with accession numbers 12897–12908.

Mouse ES Cell Differentiation and Multiple Reaction Monitoring Mass Spectrometry—For induction differentiation, the mES cells were cultured following withdrawal of LIF for various times (0, 12, and 24 h). To validate the phosphosites identified in the current study, a multiple reaction monitoring (MRM) experiment was also performed to explore the abundance profiling at different time points. The isotopic peptides (Fmoc-[13C6]Leu; labeled form) with the same sequence without iso- topo (unlabeled form) were selected, synthesized (GL Biochem, Shanghai, China), and purified by the manufacturer (>95% purity). Dried peptides were reconstituted in 0.1% formic acid in H2O (v/v) before use.

The entire analysis was performed by reversed-phase high performance liquid chromatography on a 1200 HPLC system (Agilent Technologies) coupled on line to a 6410 QQQ mass spectrometer (Agilent Technologies). In detail, the one-dimensional LC system was set up using one C18 trap column (300 µm × 5 mm, Agilent Technologies) followed by an analytical C18 column (150 µm × 100 mm, Column Technology Inc.). To elute the peptides, a gradient of 5–40% buffer B (90% acetonitrile in 0.1% formic acid) was used. The parameters for mass spectrometer were as follows: capillary voltage at 4000 V, drying gas at 300 °C at 3.0 liters/min, and nebulizer gas at 18 p.s.i. MRM transitions were optimized with MassHunter software (Agilent Technologies) and acquired at unit resolution in both the Q1 and Q3 quadrupoles to maximize sensitivity. Finally, two transitions per peptide (a quantifier and a qualifier) were selected and monitored. For quantitation analysis, 40 µg of tryptic peptides from mES cells were resublimibrated in 0.1% formic acid spiked with 1 pmol of internal standard peptides. The labeled non-phosphopeptide PSSDPLQG-
SPDSTSPK with 3+ charge state (m/z 645.6483; \text{[}^{13}\text{C}_{6}\text{]Leu}) and labeled phosphopeptide PPSETDPLLQGS'PDSTSPK' with 3+ charge state (m/z 698.9591; \text{[}^{13}\text{C}_{6}\text{]Leu}; the asterisk indicates the phosphorylated residue) were used as internal standard peptides and monitored by MRM.

For each spiked and endogenous peptide, the best transitions (non-phosphopeptide: 645.6 \rightarrow 818.4 and 643.6 \rightarrow 818.4; collision energy, 13; dwell time, 80 ms; phosphopeptide: 699.0 \rightarrow 733.8 and 697.0 \rightarrow 730.8; collision energy, 9; dwell time, 80 ms) were used for further analysis. In addition to the qualifier, the retention time was also used to confirm the peptide elution pattern. The peak area was calculated by MassHunter software (Agilent Technologies). All the results obtained from MRM experiments were averaged from triplicated measurements to decrease variation. All the MRM spectra can be found in supplemental Fig. 4.

Defining Stemness Proteins—Using previous studies, we predefined a list of 233 stemness genes, which could probably uncover the core stem cell characteristics that underlie self-renewal and generate differentiated progeny. Of these genes, 216 have been proven to be enriched in mouse embryonic, neural, and hematopoietic stem cells (38), and 15 were identified to be ES cell-specific markers (19). SALL4 (39) and REST (40) were confirmed as two main regulators in maintaining the pluripotency of stem cells. In our work, a total of 52 proteins and 22 phosphoproteins from the stemness genes were confidently identified (Table 1 and supplemental Table 5).

Protein Overlap Analysis—From three previous publications on mES cells, we extracted 846 (16), 1871 (17), and 2389 (18) proteins. Proteins with an identical sequence in each study were regarded as unique proteins, and the three studies yielded a total of 3870 unique proteins referred to in the “previous studies (3870)” data set (Fig. 2A). The same procedure was used to compare the three large data sets in Fig. 2A (previous studies (3870), previous study (5111), and our study (4581)), and 1060 shared proteins were found among them.

Phosphorylation Motif Analysis—The phosphosites of tissue cells were extracted from the “SysPTM” database. The mouse tissue phosphorylation data included mouse brain (PMID: 15345747, 18034455, and 18056256), mouse forebrain synaptosome (PMID: 15572359), mouse brain post synaptic density (PMID: 16452087), mouse cortex (PMID: 17114649), mouse forebrain (PMID: 18388127), mouse liver (PMID: 17203969 and 17242355), mouse liver and mouse heart (PMID: 18579562), mouse liver cell plasma membrane (PMID: 18266315), and mouse liver mitochondrial protein lysate (PMID: 17208939). The human tissue phosphorylation data included human brain (PMID: 18510355), human liver (PMID: 18310008 and 18491316), human T cells (PMID: 15144186), human skin fibroblast (PMID: 16774116), human mammary epithelial cells (PMID: 17389395), and human platelets (PMID: 18088087). The motif significance parameter of Motif-X (41) was set to 1e–6 (supplemental Table 6).

Conservation Analysis of Phosphoproteins and Phosphosites—Conservation of phosphoproteins between humans and mice was defined as the ratio of orthologous phosphoproteins to total phosphoproteins. The phosphoproteins and phosphosites of hES cells were extracted from published studies (22–24). The redundancy of phosphoproteins was reduced if their sequences from individual data sets were identical. The homolog gene data from mice and humans were downloaded from NCBI (HomoloGene Release 63) and were used as a reference to find the homologous and orthologous phosphoproteins between humans and mice; only one copy of multiple paralog genes in mice or humans was selected in the following analysis. As a result, the degree of phosphoprotein conservation in ES and tissue cells was 25% (1015/4019) and 17% (605/3473), respectively. A $\chi^2$ test was used to test the significance of the conservation difference of phosphoproteins between ES and tissue cells with input numbers of 1015, 4019, 605, and 3473. The resultant $p$ value was less than 1e–10. To analyze the conservation differences of phosphoproteins between ES and tissue cells in specific gene function modules (Fig. 5A), the orthologous phosphoproteins in ES cells and tissue cells were classified in gene ontology terms (biological processes and molecular functions). The significance of the conservation difference of the individual gene ontology terms was tested using the same method described above for global phosphoproteins.

Phosphosite conservation between mice and humans was only considered in the orthologous phosphoproteins of the two species: the 1015 phosphoproteins in ES cells and 605 phosphoproteins in tissue cells. The orthologous phosphosites between mice and humans were obtained by performing a BLASTP search of each orthologous phosphoprotein in ES cells and tissue cells against itself. For each orthologous phosphoprotein, the highest score of the local region of the protein sequence from the BLASTP results was used to find the orthologous phosphosites. Only one copy of the multiple paralogous phosphosites was selected from the BLASTP results in mouse or human. Conservation of phosphosites between humans and mice was defined as the ratio of orthologous phosphosites to total phosphosites. As a result, the degree of phosphosite conservation in ES and tissue cells was 29% (1924/6667) and 40% (1222/3074), respectively. A $\chi^2$ test was used to test the significance of the conservation difference of phosphosites between ES and tissue cells with input numbers of 1924, 6667, 1222, and 3074. The resultant $p$ value was less than 1e–10. The orthologous phosphosites in ES and tissue cells were classified in gene ontology terms (biological processes and molecular functions) based on the InterPro domain annotation file in ipi.HUMAN.IPC and ipi.MOUSE.IPC (ftp://ftp.ebi.ac.uk/pub/databases/ InterPro). The conservation differences of phosphosites between ES and tissue cells in specific gene function modules are shown in Fig. 5B. The significance of the conservation difference for the individual gene ontology terms was tested using the same method described above for global phosphosite analysis. The phosphosites, which were identified only in ES cells or tissue cells, were annotated in gene ontology terms and are shown in Fig. 5C with different colors. All significance analyses of specific terms of phosphosites between ES and tissue cells were conducted with a $\chi^2$ test (supplemental Table 7), and only the specific term with a $p$ value less than 0.05 was considered.

RESULTS

Experimental Strategy

The entire workflow is shown in Fig. 1. To confirm the undifferentiated state of the E14 cells used in the experiment, we first observed their morphology and performed alkaline phosphatase/SSEA-1 staining (supplemental Fig. 1). Before harvesting, the cells were examined under a microscope to ensure that the mES cells were more than 90% confluent and had a high nucleus/cytosol ratio (supplemental Fig. 1A). Immunostaining of cells cultured under the same conditions for the murine-specific stem cell surface markers alkaline phosphatase and SSEA-1 and for the expression of OCT4, SOX2, and NANOG, which are master regulators of mES cells, strongly demonstrates the undifferentiated state of the mES cells used in the assays (supplemental Fig. 1, B, C, and D).

To reduce the complexity of the mES cell proteome, we used standard cell subcellular fractionation methods (see “Experimental Procedures”) and obtained the cytosolic and nuclear extractions. The quality of the fractionation procedure
Supplemental Table 1. The data set was compared with four previous large scale studies on mES cells (16–19), and the results were categorized into three data sets (Fig. 2A). Altogether, 1060 proteins overlapped and comprised 27, 21, and 10% of the individual data sets. These independent investigations of the mES cell proteome were cross-complementary and increased the number of mES cell proteins to 9297 regardless of differences in the instruments, search engines, and final criteria used.

mES Cell Phosphoproteome Profiling

Establishment of mES Cell Phosphoproteome—To comprehensively profile the phosphoproteome of mES cells, we used not only affinity purification for phosphoproteins and SAX off-line fractionation combined with TiO₂ enrichment for phosphopeptides but also subcellular fractionation followed by chromatographic enrichment for phosphopeptides. With an FDR below 1% for peptides and phosphosites (see “Experimental Procedures”), our approach confidently identified 3970 distinct phosphosites (supplemental Table 2 and supplemental Fig. 5) distributed over 2880 unique phosphopeptides (36,682 phosphopeptides; supplemental Table 3) derived from 1642 phosphoproteins; this is the largest phosphorylation data set obtained for mES cells to date. Furthermore, to make efficient use of the data set and share it for future research (42), we developed a web site StemCell Project where users can retrieve and analyze proteome and phosphoproteome data from large scale mES cell studies.

In general, various types of phosphopeptides could be recovered as previously reported with respect to their isolation specificity (43). Using multiplex and complementary strategies, we confidently identified 1570, 1509, and 1936 unique phosphoproteins (Fig. 2B). Of the 2880 unique phosphopeptides identified, 79% were singly phosphorylated peptides, 17% were doubly phosphorylated peptides, and 4% were multiply phosphorylated peptides. Furthermore, the distribution of phosphoserine (Ser(P)), phosphothreonine (Thr(P)), and phosphorysine (Tyr(P)) was determined to be 87, 12, and 1%, respectively, which is close to the phosphorylated amino acid distribution found in hES cells (22). To archive an overview of the identified modifications, we used the PTMBlas tool in SysPTM (44) to check the status of 3970 phosphosites. In total, 1463 (37%) phosphorylation sites were found in SysPTM, and 2507 (63%) novel phosphorylation sites were identified for the first time in mES cells (Fig. 2C).

Representation of mES Cell Phosphoproteome—We investigated the phosphoproteome representation of mES cells by molecular weight, isoelectric point (pl), and amino acid composition (Fig. 2, D–F). As a comparison, the corresponding attributes of 4581 identified proteins and proteins in the IPI mouse database (version 3.52) were used as the background. The molecular weight distributions of the three data sets displayed a similar trend; however, the proteome and phosphoproteome shifted slightly toward a higher range of molecular weights. This shift could arise partly because of the limitation of the current work flow in which small proteins might be lost. The pl profiles appear as two waves representing acidic and basic proteins, and all cross in the range from pH 2 to pH 13. Because acidic phosphopeptides were more likely to be identified based on the strategy used (32), the proteins in the phosphoproteome and proteome were more likely to fall into the acidic range of the isoelectric point than those in the database. For the amino acid distribution analysis, we could not find an obvious pattern of occurrence for any amino acid appearing in the phosphoproteome, although most phosphorylation events occur at serine, threonine, and tyrosine sites. In general, the mES cell phosphoproteome
analyzed shows no obvious artificial bias and can thus be taken as a reasonable representation.

Some signaling pathways have been reported to function in maintaining the undifferentiated state of mES cells. These pathways include the JAK-STAT3, MAPK, SHP2-Ras-ERK, PI3K-AKT, mammalian target of rapamycin (mTOR), TGF, Wnt signaling pathways, and many others (45–50). Many protein components of these pathways were identified in the current work, and some were phosphorylated (supplemental Table 4).

Although enrichment and fractionation methods were applied, the relative low discovery rate for phosphorylation sites in these pathways may indicate that intense approaches for phosphorylation analysis are needed, especially for targeted research on some signaling pathways.

Kinases and phosphatases dominate the reversible conversion between the phosphorylated and dephosphorylated protein forms, which is a key process in controlling cascade signal transduction in mES cells. We identified several Ser/Thr, Tyr, and dual specificity kinase/phosphatase families (supplemental Table 4), including the CAMKs, CDKs, MAPKs, and MAP2Ks families. A total of 18 phosphorylated kinases (24 specific phosphosites) and seven phosphorylated phosphatases (12 specific phosphosites) were identified.

Most signals from the surroundings are transmitted into living cells through transmembrane proteins. We predicted the transmembrane status of the 4581 identified proteins using TMHMM2.0 (51). In total, 148 proteins were predicted to contain transmembrane helixes, and 28 of those proteins were phosphorylated (supplemental Table 4).

Identification of Stemness Phosphoproteins—To further explore the phosphosites in some key regulators of mES cells, 52 stemness proteins were chosen for analysis (supplemental Table 5). The functions of these proteins were closely associated with mES cell characteristics (see “Experimental Procedures”). Twenty-two of the proteins were found to contain specific phosphosites that had not been identified previously (Table I). It is worth noting that although some of these proteins were previously identified as being phosphorylated (NANOG (29) and REST (52)) this study was the first to confidently identify a series of specific phosphosites in these critical mES cell proteins. These proteins include NANOG (Ser-65; Fig. 3A), REST (Ser-950 and Thr-953; Fig. 3B), SALL4 (Ser-122, Ser-124, and Ser-347), RIF1 (Ser-1050, Ser-1051, Ser-1358, Ser-1461, Ser-1464, Ser-1547, Ser-1557, Ser-1572, Ser-1835, Ser-1875, Ser-2101, and Ser-2151), and UTF1 (Ser-15, Ser-18, Ser-48, Ser-54, Ser-99, Ser-171, Ser-172, and Ser-174). All of the novel phosphosites can be used to study the function and mechanism of the phosphorylation of stemness proteins in mES cells. We used the NetPhorest tool (53) to predict the putative kinases responsible for the 83
**TABLE I**

Prominent phosphorylated stem cell markers identified in current study

For all identified genes and proteins related to markers, see supplemental Table 5. For all confident identifications, the phosphorylation sites are marked with an asterisk. The literature search of the phosphorylation sites was based on the UniProt database (71). “Yes,” reported in the database; “no,” novel site. PKD, Protein kinase D; CLK, cdc2-like kinase; NEKs, NIMA related kinase family; PAKB, myosin I heavy chain kinase; DMPK, dystrophia myotonica-protein kinase; MSN, Ste20 kinase.

| IPI accession no. | Gene name | Peptide sequence | Modification site(s) | Literature | Putative kinase group |
|-------------------|-----------|------------------|----------------------|------------|-----------------------|
| IPI00421167 | Rif1 | ASQGLS*AVENS*E'S*DSSEAK | 2 | Yes, yes | CK2, CK2 |
|                 |          | LT*DES*PQENL*TEK | 2 | No, no | MARK2, CDK2,3 |
|                 |          | PAVENAS*LEDLTT*EKK | 1 | No | CK2 |
|                 |          | TYOTDASEAVSEIQGPC*SENSH*PAEDPGLSECK | 1 | No, p38 | |
|                 |          | VIAGS*S*PEGVETM*ELNR | 1 | No, p38 | |
|                 |          | SQED*E*IS*PV*NK | 1 | Yes | p38 |
|                 |          | LLEEKS*T*DFVIPPEGK | 2 | No, no | PKD, CK2 |
|                 |          | LPQCLLVQVASELGAE*SNTTS*PEK | 1 | Yes | CDK2,3 |
|                 |          | SSD*S*VDEE*QE*EKK | 1 | Yes | CK2 |
| IPI00874522 | Tjp1 | ADGATS*DDLH*HDR | 1 | No | CK2 |
|                 |          | ATLLNV*PDLS*DS*I*S'H*ANAS*ER | 4 | Yes, yes, yes, yes | ACTR2_ACTR2B_TGFbR2, NEKs, MARK2, CK2 |
|                 |          | DDISEIQSLS*DS'H*SR | 2 | Yes, yes | CK2, PKC |
|                 |          | PGS*IAST*PV*K | 1 | No | CDK2,3 |
|                 |          | S*REDLSAGPVQ*TK | 1 | Yes | CLK |
|                 |          | VQPVS*HPDPEPS*DNEDDS*Y*DEEV*H*PR | 1 | Yes | CK2 |
|                 |          | AIPVS*PS'A*VEDDEDEOD*GHTW*ATAR | 2 | Yes, yes | CDK2,3, CK2 |
|                 |          | HTSDYE*D*TD*G*AYTD*Q*DEL*T*LNDEVG*T*P*P*E*A*I*R | 1 | Yes | CDK2,3 |
| IPI00454162 | Utf1 | APVS*P*DSAQR | 1 | No | p38 |
|                 |          | GRS*S*LS*ALAPAPA*PQVEAEL*PLA*AENDEPA*PALR | 3 | No, no, no | CLK, PAKB, PKC |
|                 |          | SAGDVPV*T*DD*FAT*S*GGMA*EPGS*PK | 1 | No | CDK2,3 |
|                 |          | RVS*A*ALAR | 1 | No | DMPK |
|                 |          | RLP*AF*SPPS*P*PA*ELR | 2 | No, p38, CDK2,3, p38 | |
| IPI00318938 | Eif4ebp1 | SAGSS*C*SO*TS*PSR | 1 | No | CK2,3 |
|                 |          | VAL*DGVGQ*LPDGP*DY*T*PGGT*L*FS*TP*PGG*TR | 2 | Yes, yes | CDK2,3, MSN |
|                 |          | T**PKDL*PAI*P*V*GT*ST*S*DEPPMGAS*N*Q*SQL*P*S*EDK | 1 | No | CDK2,3 |
|                 |          | DLPAI*P*V*GT*ST*S*DEPPM*Q*A*S*N*Q*SL*P*S*EDK | 1 | Yes | CDK2,3 |
| IPI00875692 | Ercc5 | ALDDD*NEKEVS*G*SS'S*D*L*LAEK | 2 | No, no | CK2, CK2 |
|                 |          | DTEC*LQQ*DS*DE*A*M*EGHR | 1 | No | CK2 |
|                 |          | LS*S*DD*TEGQGN*NP*PK | 2 | No, no | CK2, CK2 |
|                 |          | GQDIY*VLP*PL*E*EE*K*S*E*DEDEK | 2 | No, yes | CK2, CK2 |
| IPI00123801 | Hsp1 | IES*PK*LER | 1 | Yes | p38 |
|                 |          | N IQ associated with the ES cell line. | 1 | Yes | p38 |
|                 |          | VPTEEED*GS*LE*A*DE*MECP*N*Q*RP*TE*S*D*VD*DK | 1 | Yes | CK2 |
| IPI00108388 | Usp9x | NGI*STG*S*G*SS'S*D*L*DD*MS*G*DEK | 2 | No, yes | CK2, CK2 |
|                 |          | AQ*ENYE*G*EE*V*SS*PQ*TK | 1 | Yes | CDK2,3 |
|                 |          | HGS*SEDE*E*W*LAER | 1 | No | CK2 |
| IPI00406794 | Gab1 | SSS*LE*G*F*HSQY*K | 1 | No | DMPK |
|                 |          | SV*TAGV*VGS*G*E*LE*DN*VP*K*N*P*S*N*P*PR | 1 | No | CDK2,3 |
|                 |          | SY*SH*DV*L*PK | 1 | Yes | CK2 |
| IPI00112963 | Ctnna1 | SRT*V*Q*VE*D*DDLIA*G*Q*SAR | 3 | Yes, yes, yes | DMPK, DMPK, CK2 |
|                 |          | T*PEEL*D*DS*DF*F*ED*F*VR | 2 | No, no | p38, CK2 |
specific phosphosites. The top three kinase groups were the CK2_group (casein kinase 2 for 30 phosphosites), the CDK2,3_group (CDK2 and -3 for 19 phosphosites), and the p38_group (MAPK1, -11, -13, and -14 for 15 phosphosites) (Table I). Furthermore, to validate the phospho-NANOG and -REST proteins, we also performed a dephosphorylation treatment with phosphatase. As shown in Fig. 3, C and D, the treated NANOG and REST are present as single bands on immunoblots (29, 52).

To determine the biological relevance of the novel phosphorylation sites identified in the current study, an MRM assay was developed to monitor any change in NANOG phosphorylation following LIF withdrawal at three different time points (0, 12, and 24 h). Because of the highly sensitive and robust quantitative results provided by MRM, the phosphorylated and non-phosphorylated forms of NANOG could be monitored in a single experiment (Fig. 3 E and supplemental Fig. 4). In agreement with previous reports (53) and as shown in Fig. 3 E (643.6 → 818.4), the level of NANOG decreased following LIF withdrawal. Our results clearly show that both peptide forms of NANOG exhibit decreased abundance. However, the abundance of the phosphorylated peptide (Fig. 3 E; 697.0 → 730.8) decreased rapidly immediately following LIF withdrawal (from 0 to 12 h), and no decrease could be observed from 12 to 24 h. In contrast, the abundance of the non-phosphopeptide began to fall drastically from 12 to 24 h in the absence of LIF.

**Phosphorylation Motif Discovery**

To identify significant phosphorylation motifs in ES cells compared with ultimate tissue cells, the Motif-X algorithm (41) was used to extract phosphorylation motifs from the mES cell phosphorylation data that we generated and the hES cell phosphorylation data collected from the published literature (22–24, 54). The phosphoproteins and phosphosites from mouse and human tissue cells that were used as the background data set were extracted from the SysPTM database, which stores tissue phosphorylation data collected from individual mass spectrometry-based work (see “Experimental Procedures”). To be more specific, phosphopeptides containing the 3970 confidently identified phosphosites of mES cells were prealigned at ±6 amino acids from the central phosphorylated Ser/Thr/Tyr without any extension to the end of the N or C terminus. Finally, the 3957 prealigned peptides were submitted to the Motif-X algorithm. A total of 6058 prealigned phosphopeptides of hES cells, 8927 prealigned phosphopeptides of mouse tissue cells, and 6121 prealigned phosphopeptides of human tissue cells were also submitted in parallel to the Motif-X algorithm. As a result, 19 and 26 phosphorylation motifs were discovered in mES and hES cells, respectively (supplemental Table 6). Interestingly, only Pro-directed phosphorylation SP motif was found to exist in both mouse and human ES cells (Fig. 4A). We also wondered whether the
motif existed in the identified phosphorylated master gene products. It was surprising that the pattern of the SP motif was prevalently found in some pivotal stem cell factors (NANOG, REST, SALL4, and UTF1) (Fig. 4B). The phosphosites located in the motif were potential substrates of the p38 group and CDK2,3 group (Table I). The identification of their true kinases would clarify the function of these phosphorylation events in mES cells.
Phosphorylation Conservation in ES Cells

To investigate evolutionary differences in the phosphorylation events between undifferentiated ES cells and ultimately differentiated tissue cells, we performed a conservation analysis of the phosphoproteins and phosphosites. The conservation analysis of phosphoproteins was conducted as described elsewhere (55). The data were derived from the current work on mES cells, other phosphorylation sources of hES cells, and mouse and human tissue cells (see “Experimental Procedures”).

It is important to determine whether the phosphoproteins are more conserved in the undifferentiated ES cell state or in the ultimately differentiated tissue cell state. Conservation of phosphoproteins between humans and mice was defined as the ratio of orthologous phosphoproteins to total phosphoproteins (see “Experimental Procedures”). Of the 1533 and 3501 orthologous phosphoproteins in mES and hES cells, respectively (22–24), a total of 1015 phosphoproteins were found to overlap with a degree of conservation of 25% (1015 of 4019). On the other hand, of the 2897 and 1181 orthologous phosphoproteins in mouse and human tissue cells, respectively, 605 shared phosphoproteins were found with a degree of conservation of 17% (605 of 3473). The significant \( \chi^2 \) value of the two data sets \( (p < 1.0e\text{-}10) \) indicated that, based on the present data, phosphoproteins were more conserved in ES cells than in tissue cells on a global basis (see “Experimental Procedures”). The global phosphoprotein conservation analysis described above was further separated into specific gene ontology terms (Fig. 5A and supplemental Table 7), which suggested that the biological function modules of potassium ion transport, ion channel activity, and transmembrane transport were more conserved in ES cells, whereas hydrolase activity, oxidoreductase activity, transcription, and many other activities were less conserved when compared with tissue cells.

A phosphosite was defined as conserved between humans and mice only if the orthologous sites were phosphorylated in the two species (55). The phosphosite conservation analysis was only considered in orthologous phosphoproteins of the two species (1015 phosphoproteins in ES cells and 605 phosphoproteins in tissue cells). The degree of phosphosite conservation was calculated using a method similar to that used for phosphoproteins. Of the 2995 and 5596 orthologous phosphosites in mES and hES cells, respectively, 1924 shared phosphosites were discovered with a degree of conservation of 29% (1924 of 6667). Additionally, of the 2853 and 1443 orthologous phosphosites in mouse and human tissue cells, respectively, 1222 phosphosites were found in the two
Fig. 5. Phosphorylation conservation analysis of ES cells (blue) compared with ultimately differentiated tissue cells (red). A, significantly conserved global phosphoproteins and gene ontology term-classified phosphoproteins in ES cells and tissue cells. B, significantly conserved global phosphosites and gene ontology term-classified phosphosites in tissue cells. C, functional annotation of homologous phosphosites identified between humans and mice. Error bars indicate 95% confidence interval of conservation degree.
species with a degree of conservation of 40% (1222 of 3074). Intriguingly, the conservation analysis of phosphosites yielded the opposite result observed for phosphoproteins. In general, the phosphosites in tissues were more conserved than in ES cells ($p < 1.0e^{-10}$). In tissue cells, more conserved phosphosites were found in domains possessing ATP binding and transcription regulation functions than in ES cells (Fig. 5B). However, many of the dominant phosphosites in ES cells were located in domains involved in nucleic acid binding, translation regulation, RNA processing, and DNA methylation (Fig. 5C).

DISCUSSION

In recent years, the proteome of mES cells has been widely studied; however, the phosphoproteome characteristics of mES cells have not been identified. In the current study, we integrated multiplex approaches to enrich the phosphoproteins and phosphopeptides from mES cells and confidently identified 3970 phosphosites from 1642 phosphoproteins. These data establish a large scale phosphorylation data set that contains a list of novel specific phosphosites in some key regulators of mES cells. They also allow for subsequent extensive bioinformatics analysis of the comprehensive phosphorylation features in mES cells.

The phosphorylation data that we generated included 22 phosphorylated stemness proteins with 87 specific phosphosites, most of which could be phosphorylated by the CK2, CDK2,3, or MAPK kinase group. CK2 is a positive regulator of the canonical Wnt signaling pathway, which is involved in stem cell fate determination (46, 56, 57). CDKs dominantly control the rhythm of the cell cycle and thus regulate the self-renewal of stem cells (58). Classical MAPK signaling pathways have been confirmed to regulate stem cells from the early steps of differentiation to maturation (59, 60). By phosphorylating the core stemness proteins, which induces them into active/inactive states, all of the putative kinases could orchestrate the regulation of stem cell self-renewal or differentiation (24).

NANOG is a key regulator in the maintenance of pluripotency and self-renewal in mES cells and was previously suggested to be phosphorylated, but the exact phosphosite(s) was unclear (8, 29). For the first time, we have successfully identified the exact phosphorylation sites of the NANOG protein (Fig. 3A). It is significant that the novel phosphosite(s) are located in the N terminus of NANOG, which appears to be a serine-rich sequence related to transactivation (61). Additionally, we performed a quantitative MS experiment to analyze NANOG phosphopeptide changes following cell differentiation. Unexpectedly, we observed different trends for the decrease in the phosphopeptide and non-phosphopeptide levels (Fig. 3E). Phosphatases cannot control the activity of phosphorylated proteins and protein degradation. The two different profiles may imply a unique regulatory mechanism for the phosphorylation events of NANOG. Moreover, such modifications of “master” regulators could play a role in the mechanism of pluripotency (62). However, further efforts are needed to study the biological importance and functional significance of this mechanism. Additionally, some notable proteins regulated by NANOG, including RIF1 (63) and REST (Table I), were shown to be phosphorylated in our study. For the first time, we identified the specific phosphosites in REST (Fig. 3B). These sites are a newly discovered element that is required for maintaining the self-renewal and pluripotency of mES cells (40).

The abundant SP motif in pivotal factors, including NANOG, REST, SALL4, and UTF1, may suggest that the responsible kinases are active in mES cells. It could also be hypothesized that the phosphorylation events of these prominent translation factors occur downstream of MAPKs and cell cycle-related signaling pathways. Therefore, they probably act as key hubs connecting transcriptional regulating networks and signaling pathways in mES cells, events that have been found to be necessary in maintaining the undifferentiated state of mES cells (50, 64–70).

Moreover, we studied the evolutionary difference between ES and tissue cells of human and mouse. At the global phosphoprotein level, the degree of phosphorylation conservation was higher in ES cells than in tissue cells, implying that phosphoproteins are more likely to be engaged in the essential functions of ES cells. In contrast, the degree of phosphorylation conservation was reversed at the global phosphosite level, which could suggest that the phosphosites are more variable in ES cells. The more rapid evolutionary change of the phosphosites might indicate more flexible regulatory activities in ES cells, which may contribute to adaption to new developmental processes in differentiation. As for specific biological function modules, some might be more conserved in ES cells than in tissue cells at the phosphoprotein or phosphosite level. For example, ion channel activity and transmembrane transport were more conserved in ES cells with respect to phosphoproteins (Fig. 5A). A series of phosphosites found only in ES cells were located in domains of DNA- or RNA-related processes or functions including nucleic acid binding, DNA methylation, and mRNA/rRNA processing. These biological activities have been shown to be closely connected to the maintenance of pluripotency and self-renewal in ES cells.

CONCLUSION

ES cells possess the remarkable potential to develop into all cell types, which is believed to hold great promise for regenerative medicine and cancer therapies. In the current study, we focused on the large scale analysis of phosphorylation events in mES cells at the proteomic level. We obtained the most comprehensive phosphorylation event data in mES cells to date. Altogether, 52 protein markers of mES cells were identified, including SOX2, OCT4, REST, UTF1, and NANOG. Twenty-two of those proteins were found to be phosphorylated. Additionally, we constructed a database that contains
Phosphoproteome Features of Mouse ES Cells

all of the information contained in the current study. These data extend our knowledge of the global phosphorylation state in mES cells and are a valuable resource for the understanding and future study of the basic biology of mES cells.

Acknowledgments—We thank Dr. Joshua J. Coon and Dr. Danielle L. Swaney for kindly providing the hES cell phosphoproteome data. We acknowledge Dr. Lu Xie from the Shanghai Center for Bioinformatics Technology for revising the manuscript.

This work was supported by Ministry of Science and Technology Grants 2006CB943900, 2007CB847904, 2007CB947100, 2007CB948000, 2010CB912702, 2010CB912100, and 2011CB910200; National Natural Science Foundation of China Grants 30621091 and 30821065; and Chinese Academy of Sciences Project Grant KSCX2-YW-R-106.

This article contains supplemental Figs. 1–5 and Tables 1–7.

Both authors contributed equally to this work.

**To whom correspondence may be addressed. E-mail: jhkang@sibs.ac.cn.

†† To whom correspondence may be addressed. E-mail: yxl@ sibs.ac.cn.

§§ To whom correspondence may be addressed. E-mail: zr@sibs.ac.cn.

REFERENCES

1. Evans, M. J., and Kaufman, M. H. (1981) Establishment in culture of pluripotent cells from mouse embryos. Nature 289, 154–156

2. Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., and Miyazono, K. (2007) Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. J. Cell Sci. 120, 55–65

3. Friel, R., van der Sar, S., and Mee, P. J. (2005) Embryonic stem cells: understanding their history, cell biology and signaling. Adv. Drug Deliv. Rev. 57, 1894–1903

4. Nichols, J., Zewnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, 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

5. Niwa, H., Miyazaki, J., and Smith, A. G. (2000) Quantitative expression of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113, 643–655

6. Mitsu, 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

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

8. Tanaka, T. S., Kunath, T., Kimber, W. L., Jarstad, S. A., Stagg, C. A., Usuda, M., Yokota, T., Niwa, H., Rossant, J., and Ko, M. S. (2005) Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. Genome Res. 15, 2192–2198

9. Wobus, A. M., and Boheler, K. R. (2005) Embryonic stem cell: prospects for developmental biology and cell therapy. Physiol. Rev. 85, 635–678

10. Van Hoof, D., van Oostwaard, D., Ko, M. S., Vanhoof, D., Tor, R., van der Zee, J., and Mummery, C. L. (2008) Proteomics and human embryonic stem cells. Stem Cell Res. 1, 169–182

11. Stanton, L. W., and Bakre, M. M. (2007) Genomic and proteomic characterization of embryonic stem cells. Curr. Opin. Chem. Biol. 11, 399–404

12. Unwin, R. D., Gaskell, S. J., Evans, C. A., and Whetton, A. D. (2003) The potential for proteomic definition of stem cell populations. Exp. Hematol. 31, 1147–1159

13. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. Nature 422, 198–207

14. Nagano, K., Taoka, M., Yamauchi, Y., Itagaki, C., Shinkawa, T., Nunomura, K., Okamura, N., Takahashi, N., Iizumi, T., and Isebe, T. (2005) Large-scale identification of proteins expressed in mouse embryonic stem cells. Proteomics 5, 1614–1616

15. Van Hoof, D., Passier, R., Ward-Van Oostwaard, D., Pinkse, M. W., Heck, A. J., Mummery, C. L., and Krijgsveeld, J. (2006) A quest for human and mouse embryonic stem cell-specific proteins. Mol. Cell Proteomics 5, 1261–1273

16. Williamson, A. J., Smith, D. L., Blinco, D., Unwin, R. D., Pearson, S., Wilson, C., Miller, C., Lancashire, L., Lacaud, G., Kouskoff, V., and Whetton, A. D. (2008) Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. Mol. Cell Proteomics 7, 459–472

17. Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Schöler, H., and Mann, M. (2008) Stable isotope labeling by amino acids in cell culture (SiLaC) and proteome quantification of mouse embryonic stem cells to a depth of 5,111 proteins. Mol. Cell Proteomics 7, 672–683

18. Pinkse, M. W., Blagoev, B., Gnud, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127, 635–648

19. Prudhomme, W., Daley, G. Q., Zandstra, P., and Lauffenburger, D. A. (2004) Multivariate proteomic analysis of murine embryonic stem cell self-renewal versus differentiation signaling. Proc. Natl. Acad. Sci. U.S.A. 101, 2900–2905

20. Swaney, D. L., Wagner, C. D., Thomson, J. A., and Coon, J. J. (2009) Human embryonic stem cell phosphoproteome revealed by electron transfer dissociation tandem mass spectrometry. Proc. Natl. Acad. Sci. U.S.A. 106, 995–1000

21. Prill, R. J., Xiong, W., Lee, K. B., Ciccaro, S. B., Crain, A., Xu, Y., Terskhikh, A., Snyder, E. Y., and Ding, S. (2009) Phosphoproteomic analysis of human embryonic stem cells. Cell Stem Cell 5, 204–213

22. Van Hoof, D., Muñoz, J., Braam, S. R., Pinkse, M. W., Linding, R., Heck, A. J., Mummery, C. L., and Krijgsveeld, J. (2008) Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell 5, 214–226

23. Pinkse, M. W., Mohammed, S., Gouw, J. W., van Breukelen, B., Vos, H. R., and Heck, A. J. (2008) Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in Drosophila melanogaster. J. Proteome Res. 7, 667–677

24. Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jørgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol. Cell Proteomics 4, 873–886

25. Wu, J., Shackle, Q., Liu, W., Schuller, A., and Follette, M. T. (2007) Global profiling of phosphopeptides by titania affinity enrichment. J. Proteome Res. 6, 4684–4699

26. Dai, J., Shieh, C. H., Sheng, Q. H., Zhou, H., and Zeng, R. (2005) Proteomic analysis with integrated multiple dimensional liquid chromatography/ mass spectrometry based on elution of ion exchange column using pH steps. Anal. Chem. 77, 5793–5799

27. Yates, A., and Chambers, I. (2005) The homeodomain protein Nanog and pluripotency in mouse embryonic stem cells. Biochem. Soc. Trans. 33, 1518–1521

28. Dai, J., Jin, W. H., Sheng, Q. H., Shieh, C. H., Wu, J. R., and Zeng, R. (2007) Protein phosphorylation and expression profiling by YIN-yang multidimensional liquid chromatography (YIN-yang MDLC) mass spectrometry. J. Proteome Res. 6, 250–262

29. Zhou, H., Dai, J., Sheng, Q. H., Li, R. X., Shieh, C. H., Gutman, A., and Zeng, R. (2007) A fully automated 2-D LC-MS method utilizing online continuous pH and RP gradients for global proteome analysis. Electrophoresis 28, 4311–4319

30. Dai, J., Wang, L. S., Wu, Y. B., Sheng, Q. H., Wu, J. R., Shieh, C. H., and Zeng, R. (2009) Fully automatic separation and identification of phosphopeptides by Continuous pH-gradient anion exchange online coupled with reversed-phase liquid chromatography mass spectrometry. J. Proteome Res. 8, 133–141

31. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification
Phosphoproteome Features of Mouse ES Cells

rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372
34. Perkins, D. N., Pappin, D. J., Creasy, D. M., and Cottrell, J. S. (1999) Probability-based protein identification by searching sequence data-bases using mass spectrometry data. Electrophoresis 20, 3551–3567
35. Pan, C., Kumar, C., Bohl, S., Klingmueller, U., and Mann, M. (2009) Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Mol. Cell. Proteomics 8, 443–450
36. Barsnes, H., Vizcaino, J. A., Eidhammer, I., and Martens, L. (2009) PRIDE Converter: making proteomics data-sharing easy. Nat. Biotechnol. 27, 598–599
37. Vizcaino, J. A., Cotté, R., Reutter, F., Foster, J. M., Mueller, M., Rameseder, J., Hermjakob, H., and Martens, L. (2009) A guide to the Proteomics Identifications Database proteomics data repository. Proteomics 9, 4276–4283
38. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C., and Melton, D. A. (2002) "Stemness": transcriptional profiling of embryonic and adult stem cells. Science 296, 597–600
39. Yang, J., Chai, L., Fowles, T. C., Alipio, Z., Xu, D., Fink, L. M., Ward, D. C., and Ma, Y. (2008) Genome-wide analysis reveals Sall4 to be a major regulator of pluri-potency in murine-embryonic stem cells. Proc. Natl. Acad. Sci. U.S.A. 105, 19756–19761
40. Singh, S. K., Kagalwala, M. N., Parker-Thornburg, J., Adams, H., and Majumder, S. (2008) REST maintains self-renewal and pluripotency of embryonic stem cells. Nature 453, 223–227
41. Schwartz, D., and Gygó, S. P. (2005) An iterative approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat. Biotechnol. 23, 1391–1398
42. Krijgsveld, J., Whetton, A. D., Lee, B., Lemischka, I., Oh, S., Pera, M., Mummery, C., and Heck, A. J. (2008) Proteome biology of stem cells: a new joint HUPO and ISSCR initiative. Mol. Cell. Proteomics 7, 204–205
43. Bodenmiller, B., Mueller, N. M., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat. Methods 4, 231–237
44. Li, H., Xing, X., Ding, G., Li, Q., Wang, C., Xie, L., Zeng, R., and Li, Y. (2009) SysPTM: a systematic resource for proteomic research on post-translational modifications. Mol. Cell. Proteomics 8, 1839–1849
45. Murakami, M., Ichisaka, T., Maeda, M., Oshiro, N., Hara, K., Edenhoffer, F., Kiyama, H., Yonezawa, K., and Yamanaka, S. (2004) mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. Mol. Cell. Biol. 24, 6710–6718
46. Sato, N., Meijer, L., Skåltsoons, L., Greengard, P., and Brivanlou, A. H. (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat. Med. 10, 55–63
47. Watanabe, S., Umehara, H., Murayama, K., Okabe, M., Kimura, T., and Nakano, T. (2006) Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. Oncogene 25, 2697–2707
48. Davey, R. E., Onishi, K., Mahdavi, A., and Zandstra, P. W. (2007) Life-mediated control of embryonic stem cell self-renewal emerges due to an autoregulatory loop. FASEB J. 21, 2020–2032
49. Paling, N. R., Wheeldon, H., Bone, H. K., and Welham, M. J. (2004) Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling. J. Biol. Chem. 279, 48063–48070
50. Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998) Self-renewal of embryonic stem cells and embryo-derived pluripotent stem cells. Nat. Genet. 10, 921–926
51. Birney, E., Heasley, L., Bost, F., Caron, L., and Audidi, M. (2007) Concise review: regulation of embryonic stem cell lineage commitment by mitogen-activated protein kinases. Stem Cells 25, 1099–1065
52. Chen, Y., Duz, Z., and Yao, Z. (2006) Roles of the Nanog protein in murine F9 embryonal carcinoma cells and their endoderm-differentiated counterparts. Cell Res. 16, 641–650
53. Buhr, N., Carapito, C., Schaaf, C., Kieffer, E., Van Dorsseelaer, A., and Vivielle, S. (2008) Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. Electrophoresis 29, 2381–2390
54. Adams, I. R., and McLaren, A. (2004) Identification and characterisation of mRIf1: a mouse telomere-associated protein highly expressed in germ cells and embryo-derived pluripotent stem cells. Dev. Dyn. 229, 733–744
55. Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C. L., Ruan, Y., Lim, B., and Ng, H. H. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. 38, 431–440
56. Pan, G., and Thomson, J. A. (2007) Nanog and transcriptional networks in embryonic stem cell pluripotency. Cell Res. 17, 42–49
57. Darnell, E. J., Jr., Kerr, I. M., and Stark, G. R. (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science 264, 1415–1421
58. Burdon, T., Stacey, C., Chambers, I., Nichols, J., and Smith, A. (1999) Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev. Biol. 210, 30–43
59. Takahashi, K., Mitsuishi, K., and Yamanaka, S. (2003) Role of ERα in promoting tumour-like properties in mouse embryonic stem cells. Nature 423, 541–545
60. Takahashi, K., Murakami, M., and Yamanaka, S. (2005) Role of the phosphoinositide 3-kinase pathway in mouse embryonic stem (ES) cells. Biochem. Soc. Trans. 33, 1522–1525
61. Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999) STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. 18, 4261–4269
62. UniProt Consortium (2009) The Universal Protein Resource (UniProt) 2009. Nucleic Acids Res. 37, D169–D174

In order to cite this article properly, please include all of the following information: Li, Q.-R., Xing, X.-B., Chen, T.-T., Li, R.-X., Dai, J., Sheng, Q.-H., Xin, S.-M., Zhu, L.-L., Jin, Y., Pei, G., Kang, J.-H., Li, Y.-X., and Zeng, R. (2011) Large Scale Phosphoproteome Profiles Comprehensive Features of Mouse Embryonic Stem Cells. Mol. Cell. Proteomics 10(4):M110.001750. DOI: 10.1074/mcp.M110.001750.