

Research article

Open Access

Proteomic-based identification of maternal proteins in mature mouse oocytes

Ping Zhang, Xiaojian Ni, Ying Guo, Xuejiang Guo, Yufeng Wang, Zuomin Zhou, Ran Huo* and Jiahao Sha

Address: Laboratory of Reproductive Medicine, Department of Histology and Embryology, Nanjing Medical University, Nanjing, PR China

Email: Ping Zhang - shping1216@tom.com; Xiaojian Ni - nixiaojian02010630@126.com; Ying Guo - guoguo19830330@163.com; Xuejiang Guo - guo_xuejiang@njmu.edu.cn; Yufeng Wang - griffithmail@tom.com; Zuomin Zhou - zhouzm@njmu.edu.cn; Ran Huo* - huoran@njmu.edu.cn; Jiahao Sha - shajh@njmu.edu.cn

* Corresponding author

Published: 3 August 2009

Received: 24 February 2009

BMC Genomics 2009, 10:348 doi:10.1186/1471-2164-10-348

Accepted: 3 August 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/348>

© 2009 Zhang et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The mature mouse oocyte contains the full complement of maternal proteins required for fertilization, reprogramming, zygotic gene activation (ZGA), and the early stages of embryogenesis. However, due to limitations of traditional proteomics strategies, only a few abundantly expressed proteins have yet been identified. Our laboratory applied a more effective strategy: one-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (1D SDS-PAGE) and reverse-phase liquid chromatography tandem mass spectrometry (RP-LC-MS/MS) were employed to analyze the mature oocyte proteome in depth.

Results: Using this high-performance proteomic approach, we successfully identified 625 different proteins from 2700 mature mouse oocytes lacking zona pellucidae. This is the largest catalog of mature mouse oocyte proteins compiled to date. According to their pattern of expression, we screened 76 maternal proteins with high levels of mRNA expression both in oocytes and fertilized eggs. Many well-known maternal effect proteins were included in this subset, including MATER and NPM2. In addition, our mouse oocyte proteome was compared with a recently published mouse embryonic stem cell (ESC) proteome and 371 overlapping proteins were identified.

Conclusion: This proteomics analysis will be a valuable resource to aid in the characterization of important maternal proteins involved in oogenesis, fertilization, early embryonic development and in revealing their mechanisms of action.

Background

Mammalian reproduction is a complicated physiological process involving many important events, such as generation of mature gametes, fertilization, zygotic gene activation (ZGA), and embryonic development. Thus far, the key molecules and mechanisms involved in these events remain poorly characterized. Mammalian oocytes, a

highly specialized cell type, play unique roles in reproduction because only in these cells are maternal proteins and transcripts crucial for the above-mentioned processes.

During oogenesis, oocytes synthesize and accumulate a number of maternal proteins. Some of them function in the formation of follicles and/or the growth of the

oocytes, including, Fig α , GDF9, and BMP15 [1-3]. However, many maternal proteins stored in oocytes play significant roles in later stages, namely fertilization and early embryogenesis. The corresponding genes are called maternal effect genes [4], and we call the proteins they code for maternal effect proteins. Maternal effect genes/proteins have been shown to be important in early embryonic development of *Drosophila melanogaster* and *Xenopus laevis* [5,6]. Several maternal-effect genes/proteins have recently been identified in mammals, and their importance in embryonic development has also been demonstrated. MATER (Maternal antigen that embryos require; official name Nlrp5) is one of the first characterized maternal effect proteins in mice, the absence of which precludes embryonic progression beyond the 2-cell stage [7]. Npm2 is another well characterized maternal effect protein, which is required for nuclear and nucleolar organization during embryonic development [8]. Much research has been done to identify maternal effect genes or proteins essential for preimplantation or postimplantation mouse embryo development. Dppa3, Padi6, Tle6 and Floped were successfully identified in individual studies [9-11], but there remain many unknown players. Therefore, the identification and molecular characterization of novel maternal proteins will be of great significance and novel proteomic technologies can potentially deduce most of the maternal proteins in mature oocytes.

There are several recent reports utilizing proteomics approaches to the study of oocytes, including the exploration of the bovine, pig and mouse oocyte proteomes [12-16]. For example, Calvert et al. identified 8 highly abundant heat shock proteins (HSPs) and related chaperones in the mature mouse egg by two-dimensional electrophoresis (2DE) [15]. Vitale et al. used 2DE and mass spectrometry (MS) to identify 12 proteins that appeared to be differentially expressed between germinal vesicle (GV) and metaphase II (MII) murine oocytes [16]. In our previous work that demonstrated post-translational modifications of maternal proteins, we used a similar approach to perform large-scale protein identification in mature mouse oocytes, and we successfully identified a total of 380 different proteins corresponding to 869 protein spots [17]. The 2DE platform is valuable to analyze heterogeneity of proteins in the forms of alternative splicing, post-translation modifications, etc [18,19]. Although 2DE continues to be a very popular tool for studying the proteome, it has some limitations in identifying proteins that have either high or low molecular masses, those with extreme isoelectric points (pIs), those that are highly hydrophobic, and those of low abundance [20]. 1D SDS-PAGE liquid chromatography tandem mass spectrometry (LC-MS/MS) a combination of 1DE protein separation and LC-MS/MS analysis has been used widely and is generally accepted as a more effective method of studying the pro-

teome [21,22]. It is technically simple and combines improved protein separation capability that also captures those proteins typically not accessible via 2D PAGE (notably large proteins and those with transmembrane domains) with the well-established sensitivity of gel-based protein identification using MS for less complex samples [23]. For the purpose of identifying novel maternal proteins, we employed this high-performance proteomic approach to analyze proteins extracted from 2700 mature mouse oocytes lacking zona pellucida and we successfully identified 625 different proteins. The maternal protein compilation provided here is intended to serve as an important tool for expanding our knowledge of the regulation of multiple processes in mammalian reproduction.

Results

Identification of Mature Oocyte Proteins

The mouse oocyte zona pellucida (ZP) is a thick extracellular coat containing, on average, ~3.5 ng of glycoprotein which contributes approximately 15% of the total egg protein [24]. Given that high-abundance proteins may interfere with the identification of other proteins, the ZP was removed by treating with acid Tyrode solution [25]. In total, 2700 ZP-free MII oocytes were collected and the integrity of the oocytes was checked rigorously according to the criteria outlined in the Materials and Methods to ensure only morphologically normal oocytes were chosen for further research (Figure 1). Oocytes passing these selection criteria were lysed and their proteins were separated by 1D SDS-PAGE. The gel was then cut into 29 slices, proteins were in-gel digested with trypsin, and the resulting peptides extracted from each gel slice were analyzed by automated reverse phase LC (RP-LC) coupled with MS/MS. From 29 LC-MS/MS runs, a total of 42711 MS/MS spectra were acquired and searched against the IPI Mouse database v3.30 by SEQUEST. To experimentally verify the false discovery rate (FDR) in our dataset, all output files were searched against the reversed IPI Mouse database, yielding an FDR of ~1.8%. We applied additional filter criteria to exclude proteins identified with low probabilities. The confidence score of the protein was calculated by applying the PeptideProphet algorithm using Scaffold software (v01_07_00; Proteome Software) [26], and only proteins with confidence scores of more than 90% were included in our dataset. Identical peptide sequences are occasionally shared by biologically distinct proteins, thus it is occasionally difficult to determine the identities of proteins based on sequenced peptides unless unique peptides are identified. In our analysis, proteins with shared peptides were organized into a single group (protein group) in Scaffold. If a protein group comprised only isoforms or overlapping database entries indistinguishable by MS/MS analysis, then the proteins in this group were counted as a single protein. If these proteins were the

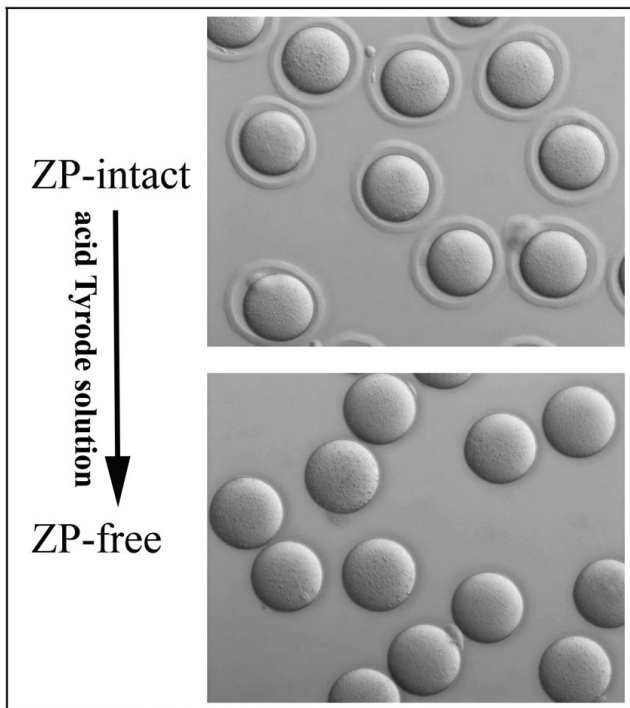


Figure 1
A representative image of a ZP-free MII oocyte selected for proteomic analysis. During the process of removing zona pellucidae, the status of oocyte was checked rigorously, and only healthy denuded oocytes were chosen for further study.

products of distinct genes, then all the proteins in the group were discarded from our dataset. Consequently, the final number of identified proteins was lower than the actual number of proteins in the sample. In summary, our dataset included 625 proteins corresponding to 611 known genes and 11 proteins from uncharacterized genes [see Additional file 1]. MS/MS spectra and fragment assignments of single peptide-based identifications are provided [see Additional files 2, 3 and 4].

We were interested to compare the mature oocyte proteins identified in the present study with those discovered by other analyses, including the 2DE study conducted in our lab as well as two other published reports [15-17]. Protein IDs in the datasets from each study were converted to gene symbols. As shown in Figure 2, a sum of 369 unique gene products were reported previously. Of these, 216 (67.7%) were also found in our present dataset, whereas 395 gene products were found only in our present study. Therefore a total of 764 different maternal proteins have been identified from mature mouse oocytes thus far.

GO and Pathway Analysis of the Identified Proteins

In Figure 3 we have the categorized proteins identified in this study in terms of cellular components based on Gene

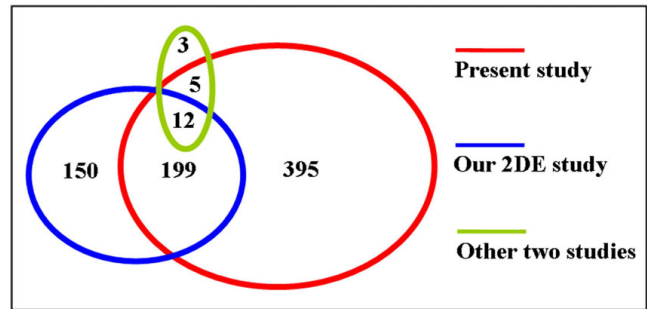


Figure 2
Diagram of proteins identified in different studies. Maternal proteins identified in the present study were compared with three recently published proteomics datasets, including the 2DE study conducted in our lab [15] and two other published reports [16,17]. Numbers represent the number of shared proteins in the respective overlapping area.

Ontology (GO) analysis. A majority of proteins (393) were assigned to the cytoplasmic compartment, accounting for 35% of the identified proteins. Among the proteins classified by GO annotation, 16% (178) were membrane proteins, followed by proteins of unknown localization (12%), nuclear (11%), and mitochondrial (5%) proteins.

To examine what biologically important entries are enriched in the mouse oocyte, we compared our dataset with a combined database of multiple mouse tissues as a reference. Generation of this combined proteomic database was based on the available proteomic data from different tissues or cells created with the same or similar LC-MS/MS experimental approach. This strategy can overcome the biases caused by the experimental approach [27]. As a result, ~28,600 genes with duplicates were included in the

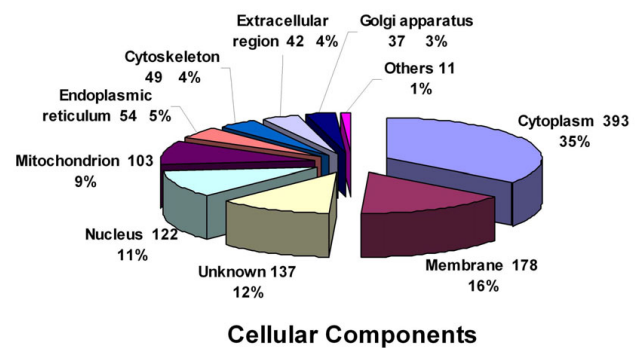


Figure 3
Distribution of oocyte proteins among subcellular compartments. The classification of the 625 identified maternal proteins was performed according to the gene ontology term "Cellular component".

combined dataset. Babelomics <http://babelomics.bioinfo.cipf.es/EntryPoint?loadForm=fatigo>, an advanced set of tools for the functional profiling of high-throughput transcriptomic, genomic and proteomic data [28], was used to carry out the calculation. When compared to the pooled proteomic database, 3 GO terms (unfolded protein binding, oxidoreductase activity, acting on CH-OH group of donors, GTPase activity) for molecular function appeared to be significantly overrepresented, and none significantly underrepresented. In the biological process category, 3 GO terms (RNA metabolic process, transcription, regulation of cellular metabolic process) appeared to be significantly underrepresented, and none significantly overrepresented (Figure 4). We found that transcription in the mouse oocyte was also underrepresented. Recent studies have demonstrated that oocytes undergo large-scale chromatin modifications in the process of maturation, including acetylation and methylation of the histone proteins, and finally global transcriptional repression appears [29]. So, the global transcriptional activity of the MII oocytes is very low, and this is consistent with our results.

A key function of maternal proteins accumulated in mature oocyte is the regulation of early embryogenesis. A detailed analysis of embryonic development influenced by the protein profile of the mouse oocyte was performed using Pathway Studio, an automated text-mining tool which enables the software to generate pathways from entries in the PubMed database as well as other public sources. Pathway analysis revealed that 65 (11%) unique gene products were involved in early embryogenesis events (Figure 5).

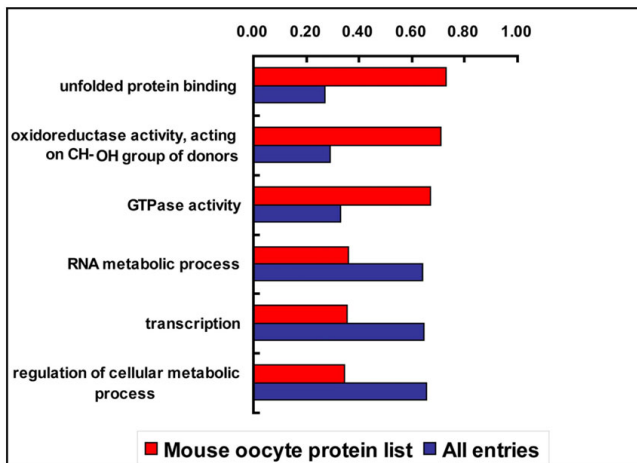


Figure 4
Over- and underrepresented annotations of the set of identified maternal proteins. The set of identified maternal proteins was compared with the combined database of multiple mouse tissues, and significantly GO molecular function and biological process terms ($p < 0,05$) are illustrated.

Analysis of High Abundance Proteins

The number of unique peptides identified is generally accepted as a semiquantitative measure of protein abundance [30,31]. We sorted the 625 proteins in our dataset according to the number of unique peptides identified. The highest ranked 23 proteins represented by more than 10 unique peptides are listed [see Additional file 5]. Among the high abundance proteins, some members have already been well characterized in oocytes. For example, DNA methyltransferase 1 (Dnmt1), is the predominant form of DNA (cytosine-5-)-methyltransferase in mammals and is essential for embryo development [32]. This protein was represented by 25 unique peptides and was the second-most abundant protein in our dataset.

Screening for Proteins with Special mRNA Expression Patterns

Unique or atypical mRNA expression patterns of maternal proteins suggest their key roles in oogenesis or early embryo development and this notion has been validated by other functional research on maternal effect proteins [7-11]. We adopted an *in silico* approach to investigate the expression patterns of proteins in our dataset by examining their expression in the mouse transcriptome database SymAtlas <http://symatlas.gnf.org/SymAtlas/>, which describes gene expression patterns in 45 mouse tissues [33]. Surprisingly, many of the proteins in our dataset exhibited patterns worthy of note. We were particularly interested in genes highly expressed in oocytes and fertilized eggs (where expression was 10-fold higher than the corresponding median values). In total, 76 proteins, representing approximately 12% of the entire dataset, were included in this subset [see Additional file 5]. As expected, many well known maternal effect proteins in mice were found in this subset. Included in this subset were the earliest identified maternal protein MATER, as well as the recently identified TLE6 [11]. In addition to these well-known maternal effect proteins, many of the proteins we identified as highly expressed in oocytes and fertilized eggs have not been characterized.

To explore the functions of the 76 proteins with abundant mRNA expression levels, we analyzed their domain composition with Babelomics. Statistical analysis indicated that the most significantly overrepresented domain was the F-box domain (Pfam: 6 proteins, $p = 3.9E-05$) and the 6 maternal proteins containing F-box domains are listed [see Additional file 5]. Different F-box proteins, as components of the Skp1-Cullin1-F-box (SCF) complex, can recruit particular substrates for ubiquitination and play central roles in cell-cycle regulation [34].

Intersection Between the Oocyte and ESC Proteome

Somatic cells can be reprogrammed by transferring their nuclear contents into oocytes [35] or by fusion with

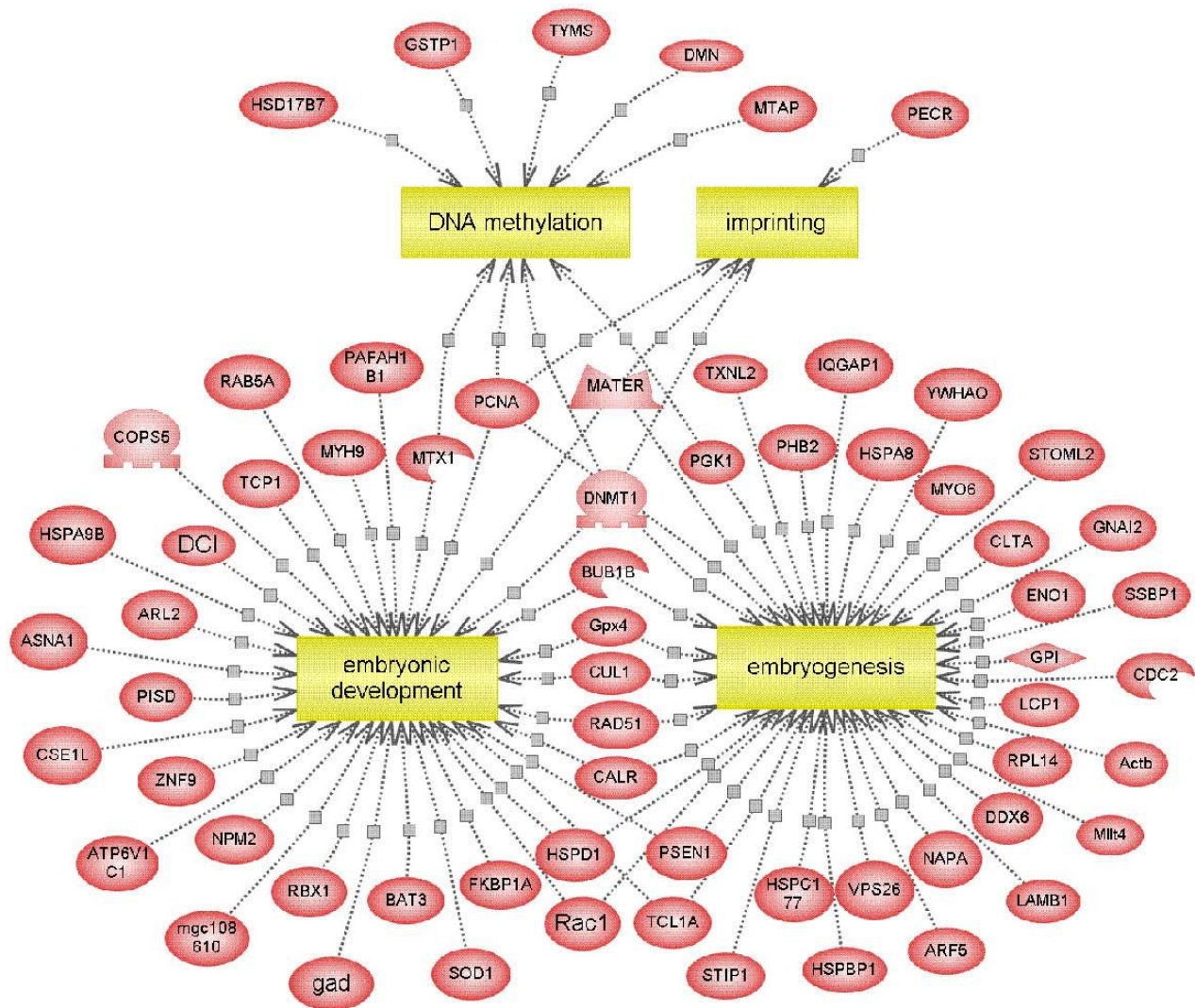


Figure 5
Biological pathways deduced for the identified maternal proteins using Pathway Studio software. Proteins involved in early embryogenesis events are shown. Proteins are indicated as red ovals and regulated processes are represented by yellow squares. Regulation events are displayed with arrows and documented by literature citations.

embryonic stem (ES) cells [36,37]. This fact indicates that unfertilized eggs and ES cells probably contain similar factors that can confer totipotency or pluripotency to somatic cells. In order to identify a common set of proteins shared between oocytes and ES cells, mouse MII oocyte proteins were compared with recently published data for proteins expressed in mouse ES cells [31]. This analysis resulted in an overlap of 371 proteins [see Additional file 5]. As expected, developmental pluripotency associated 5 (DPPA5), which has been implicated in cell pluripotency, was included in this set. Of the 371 proteins, 108 (29%) were either exclusively found in mouse ES cells or highly

enriched in ES cells compared to differentiated ES cells [see Additional file 5] [31]. It is likely that novel factors associated with reprogramming are included in this subset.

Discussion

Activation of the embryonic genome in mice begins late in one-cell zygote and is fully underway by the two-cell cleavage stage [38]. The silencing of nuclear transcription occurring between meiotic maturation in oocytes and activation of the embryonic genome implies critical roles for preexisting stores of proteins and transcripts [39].

Through knockout and knockdown strategies, individual maternal proteins have been demonstrated as essential for cleavage stage development in mice. In the present study, we identified many new and unknown maternal proteins in mice by constructing an MII oocyte proteome. In-depth analysis of these maternal proteins will assist us in screening for a proportion of great interest.

Interestingly, many maternal transcripts deposited in mammalian oocytes are not polyadenylated and therefore not translated into proteins [40]. Independent confirmation of the protein expression of maternal genes is therefore necessary. This was also an important reason for us to construct the oocyte proteome. As a case in point, NALP14, NALP5 (MATER), and NALP4f were included in our subset of abundantly expressed maternal proteins. These three proteins belong to the multifunctional NACHT nucleoside triphosphatase (NTPase) family. NALP14 and NALP5 were previously reported as maternal effect proteins and play significant roles in mouse preimplantation embryo development [7,41,42]. NALP4f was represented by 14 unique peptides in our present study and a previous analysis demonstrated that *NALP4f* was an oocyte-specific gene [42]. Our research has independently confirmed the high protein expression level of NALP4f in mature oocytes. Assuming that NALP4f has similar roles to NALP14 and NALP5, it is highly likely that NALP4f is an important factor necessary for normal embryogenesis and is a good candidate to be a maternal effect protein. In addition, we identified NALP2, NALP4b, and NALP9b in our oocyte proteome. Although the precise functions of

NACHT NTPase family members remain to be determined, we speculate that these members play significant roles in early embryo development based on their homology to NALP14 and NALP5. A distinguishing characteristic of maternal effect proteins identified to date is that the majority of them have an abundant mRNA expression in oocytes and many are expressed only in oocytes [7-11]. This fact led us to filter maternal products in our protein list by analyzing their corresponding mRNA expression patterns. As a result, 76 maternal proteins with high mRNA expression levels in oocytes and fertilized eggs were selected out. Of these proteins, we discovered that 9 previously described maternal effect proteins (MATER, STELLA, DNMT1, ZAR1, NPM2, PADI6, TLE6, TCL1, FILIA) were enriched in this subset. These maternal effect proteins have been reported to be absolutely necessary for oogenesis, fertilization or early embryo development. Indeed, apart from these well-known examples, the majority of proteins in this subset have not been previously studied or reported in oocytes. We suggest that these proteins are excellent candidates as maternal effect proteins.

A group of proteins belonging to the T-cell leukemia/lymphoma 1 (TCL1) protein family was of particular interest because their corresponding genes had dramatically similar mRNA expression patterns. Figure 6 demonstrates that *TCL1*, *TCLB1*, *TCLB2* are almost oocyte-specific genes. *TCL1* was initially identified as a gene involved in recurrent chromosomal translocation in human polymorphic leukemia (T-PLL) and overexpression of *TCL1*

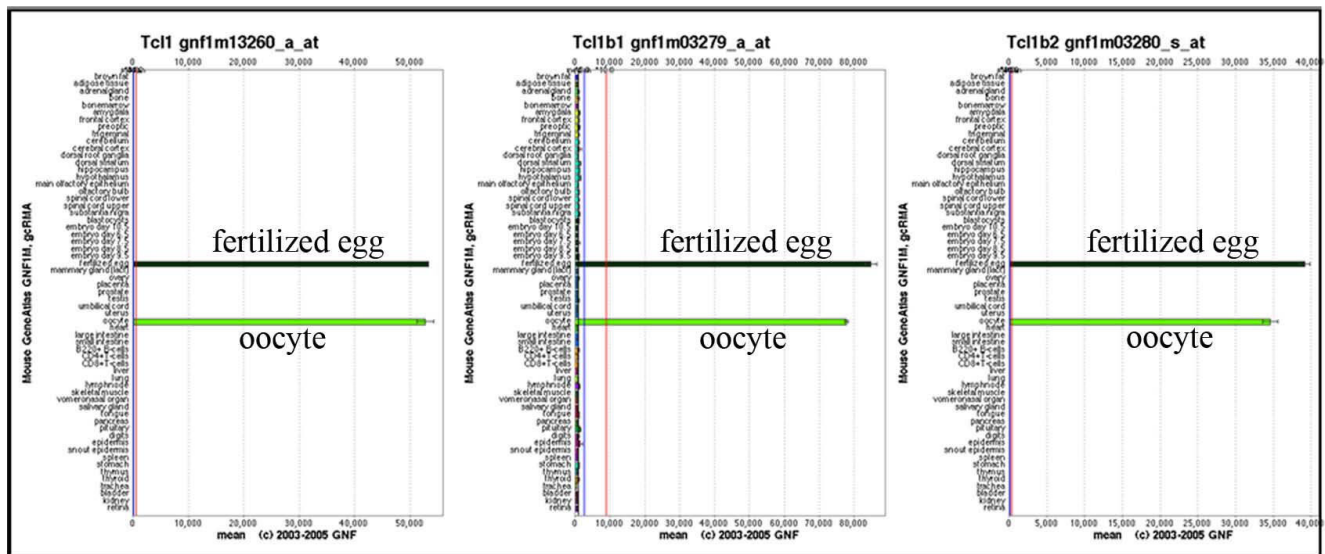


Figure 6
mRNA expression patterns of three maternal proteins. GNF SymAtlas database analysis shows that *TCL1*, *TCLB1* and *TCLB2* are almost solely expressed in the mouse oocyte and fertilized egg.

played a causative role in T cell leukemias of humans and mice [43]. However, in *TCL1*-deficient mice, a female fertility defect was observed. *TCL1*-deficient females display normal oogenesis and rates of oocyte maturation/ovulation and fertilization, but the lack of maternally derived *TCL1* impairs the embryo's ability to undergo normal cleavage and develop to the morula stage, especially under *in vitro* culture conditions [44]. The *TCL1* loss-of-function phenotype indicates that maternal protein *TCL1* plays a significant role in early embryo development. Unfortunately the functions of *TCLB1* and *TCLB2* have not yet been investigated and we can speculate that the two proteins may play similar or complementary roles in embryogenesis.

Domain composition analysis is an effective way to predict the functions of proteins identified in a proteomics analysis. Among 76 proteins singled out because their corresponding genes are highly expressed in oocytes, 6 proteins (*FBXL10*, *FBXW14*, *FBXW16*, *FBXW19*, *EG382106*, *E330009P21Rik*) contained an F-box domain, which was first described as a sequence motif in cyclin-F that interacts with the *SKP1* protein. Different F-box proteins, as substrate-specific adaptor subunits of the *Skp1-Cullin1-F-box (SCF)* complexes, recruit particular substrates for ubiquitination via specific protein-protein interaction domains. Coincidentally, three core protein subunits (*SKP1*, *RBX1*, *CUL1*) of the *SCF* complex were all definitively identified in our proteome. As one of the major classes of ubiquitin ligases, the *SCF* complex plays a central role in cell-cycle regulation [34]. In early stages of embryo development, degradation of maternal proteins is crucial for the oocyte-to-embryo transition [45]. Our results suggest that the maternal *SCF* complex probably exists in oocytes and may be important for the oocyte-to-embryo transition by recruiting specific substrates for degradation.

Pluripotent stem cells are of considerable current interest as they can proliferate indefinitely *in vitro* and give rise to many adult cell types, serving as a potentially unlimited source for tissue replacement in regenerative medicine. Recently, Takahashi et al. demonstrated that pluripotent stem cells can be induced from mouse fibroblasts by retroviral introduction of *Oct3/4*, *Sox2*, *c-Myc* and *Klf4* [46], indicating that the combination of these four factors can induce reprogramming of somatic cells to a pluripotent state. However, the use of retrovirus-transduced oncogenes represents a serious barrier to the eventual use of reprogrammed cells for therapeutic application because of tumor formation by *c-myc* reactivation [47]. Therefore it is necessary to discover factors responsible for reprogramming that would be safer for therapeutic use. We compared the maternal proteins in our oocyte proteome with a recently published mouse ES cell proteome and identi-

fied an overlap of 371 proteins. In addition to some pluripotency markers, this group included many uncharacterized proteins, some of which may be good candidates for studying the mechanism of reprogramming. A good example is translationally-controlled tumor protein (*Tpt1*), which facilitates the first step of somatic cell reprogramming [48]. Recent studies on *Tpt1* demonstrate that this protein activates transcription of *oct4* and *nanog* in transplanted somatic nuclei [49]. We believe that further analysis of these candidate proteins at the functional level will uncover novel proteins that are essential for reprogramming and indirectly promote the application of therapeutic cloning.

Conclusion

In this study, we used 1D SDS-PAGE and RP-LC-MS/MS to investigate the maternal proteins stored in mature mouse oocytes. This high-performance strategy allowed us to define a set of 625 different mouse MII oocyte proteins. This is the largest catalog of mature mouse oocyte proteins compiled to date. We believe that this study will help us to understand the diverse biological processes occurring in mouse oocytes and during early embryo development. However, compared with proteomic analyses of other cells and tissues, such as embryonic stem cells and liver, the proteins identified in mature mouse oocytes were limited. This was mainly because of the fact that mature oocytes obtained from each mouse were very limited. We believe that the catalog of maternal proteins present in this article is a starting point and we anticipate that more research on the oocyte proteome will deduce most of the maternal proteins.

Methods

All experiments requiring the use of animals received prior approval from Nanjing Medical University and were performed according to USDA-approved protocols.

Reagents

Urea (Cat. No. 17-1319-01), 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) (Cat. No. 17-1314-01), iodoacetamide (Cat. No. RPN6302), and Dithiothreitol (DTT) (Cat. No.17-1318-02) were from GE Healthcare (Uppsala, Sweden); Thiourea (Cat. No. T7875), acetonitrile (ACN) (Cat. No. 34851), ammonium bicarbonate (NH_4HCO_3) (Cat. No. A6141), and trifluoroacetic acid (TFA) (Cat. No. T0274) were from Sigma Chemical (St. Louis, MO); Protease inhibitor cocktail (Cat. No.78437) was purchased from Pierce Biotechnology (Rockford, IL).

Oocyte collection and protein extraction

Mature oocytes were obtained from ICR female mice weighing 2530 g. The mice were superovulated by intraperitoneal injection of 10 IU pregnant mare serum gona-

dotropin followed by 10 IU human chorionic gonadotropin after 48 h. After 1416 h, oocyte-cumulus cell complexes were collected from the ampulla of the oviduct, and the cumulus cells were removed by brief exposure to 1 mg/ml hyaluronidase (Sigma Chemical, St. Louis, MO, USA). Zona pellucidae (ZP) were removed by treating oocytes for a few seconds with acid Tyrode solution (pH 2.5) followed by mechanical shearing. During this process, oocyte morphology was monitored rigorously and the oocytes with shape abnormalities or with cytoplasmic abnormalities (dark cytoplasm, granular cytoplasm, and refractile body) were discarded. Only the denuded oocytes with normal morphology were selected for further investigation. ZP-free oocytes were washed 3 times in 0.01 M PBS, and stored in lysis buffer at -80°C until needed. The lysis buffer consisted of 7 M urea, 2 M thiourea, 4% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 65 mM dithiothreitol (DTT), and 1% (v/v) protease inhibitor cocktail.

One-dimensional SDS-PAGE and in-gel digestion

In brief, proteins extracted from the 2700 MII oocytes were dissolved in SDS-PAGE loading buffer, boiled for 5 min, and loaded in a single lane on a 1-mm-thick 10% polyacrylamide gel. After separation, the gel was visualized by silver staining according to a published procedure [50], except that glutaraldehyde was omitted in the sensitizing solution. Thereafter, the gel was cut into 29 slices, and each slice was cut into 1-mm³ gel particles for in-gel digestion. In-gel digestion was performed as follows: gel particles were washed 3 times in deionized water and subsequently dehydrated with 100% acetonitrile (ACN) for 10 min. The particles were incubated with 10 mM DTT in 25 mM ammonium bicarbonate for 1 h at 56°C for protein reduction. The resulting free thiol (-SH) groups were subsequently alkylated by incubating the samples with 55 mM iodoacetamide in 25 mM ammonium bicarbonate for 45 min in the dark. Gels were washed with 25 mM ammonium bicarbonate and 50% ACN solution and dehydrated with 100% ACN sequentially. The gel pieces were rehydrated with 10 ng/μl trypsin (Promega, Madison, WI, USA) in 25 mM ammonium bicarbonate and incubated for 12 h at 37°C for protein digestion. Supernatants were transferred to fresh tubes, and the remaining peptides were extracted by incubating the gel pieces twice with 30% ACN in 3% trifluoroacetic acid (TFA), followed by dehydration with 100% ACN. The extracts were combined and lyophilized to dryness, and the resulting peptides were used for mass spectrometric analysis.

Online reverse-phase LC-MS/MS

For capillary reverse-phase LC (cLC) and mass spectrometric analysis, 29 fractions were sequentially loaded onto a Michrom peptide CapTrap (MW 0.550 kD, 0.5 × 2 mm; Michrom BioResources, Inc., Auburn, CA) at a flow

rate of 50 μl/min with buffer A (see below). The trap column effluent was then transferred to a reverse-phase microcapillary column (0.1 × 150 mm, packed with Magic C18, 5 μm, 100 Å; Michrom Bioresources, Auburn, CA). The reverse-phase separation of peptides was performed using the following buffers: 5% ACN, 0.1% formic acid (buffer A) and 95% ACN, 0.1% formic acid (buffer B); a 56-min gradient (54.5% buffer B for 41 min, 90% buffer B for 5 min, and 5% buffer B for 10 min) was used. Peptide analysis was performed using Finnigan LTQ ORBITrap (ThermoFinnigan, San Jose, CA) coupled directly to an LC column. An MS survey scan was obtained for the m/z range 4001800, and MS/MS spectra were acquired from the survey scan for the 10 most intense ions (as determined by Xcaliber mass spectrometer software in real time). Dynamic mass exclusion windows of 60 s were used, and siloxane (m/z 445.120025) was used as an internal standard.

Database search and bioinformatics

DTA files (Bioworks version 3.3) in ASCII format for each MS/MS spectrum with a minimum ion count of 8 were generated from the raw data for the peptide mass range of 4008,000. The resulting spectra were independently searched against the International Protein Index Mouse database (ipi.MOUSE.v3.30, downloaded from <http://ftp.ebi.ac.uk/pub/databases/IPI>) containing 56450 entries by using SEQUEST analysis software (Bioworks version 3.3, ThermoFinnigan). Carbamidomethylation of cysteine was set as a fixed modification, and oxidized methionine was sought as a variable modification. The initial mass tolerances for protein identification on MS and MS/MS peaks were 10 ppm and 0.6 Da, respectively. Two missed cleavages were permitted. The criteria used for filtering peptides with low confidence scores were the following: cross-correlation values (Xcorr) greater than 2.0 and 2.5 were used for doubly charged ions and triply or higher charged ions, respectively; ΔCn values (difference in Xcorr with the next highest value) less than 0.1 were removed from the matched sequences. Singly charged ions were discarded because they were small in number. All output files were searched against the forward and reversed IPI mouse database separately, and FDR for all peptide-to-spectrum matches was calculated as $FDR = \frac{\# \text{ of False peptides}}{\# \text{ of True peptides} + \# \text{ of False peptides}}$.

For bioinformatics analysis, each IPI accession number was converted to an Entrez Gene ID according to the IPI protein cross-references file downloaded from <http://ftp.ebi.ac.uk/pub/databases/IPI>. We used Babelomics to find statistically over- and underrepresented GO categories in our oocyte proteome dataset. To compare the mouse oocyte proteome with other mouse tissue proteomes, we generated a combined database for the mouse based on the following mouse tissues and cell cultures

characterized by LC-MS/MS: mouse heart [51], liver [51-53], brain [51,54], lung [51,55], kidney [51], spleen [56], placenta[51], cortical neurons cell culture[57], sperm [58], islet alpha-cell culture [59]. For enrichment analysis, our identified oocyte proteome was set as a test dataset and the combined mouse proteome was set as a reference. The enrichment analysis was done using 'fisher exact test', and all GO terms that were significant with adjusted $P < 0.05$ (after correcting for multiple term testing by using the FDR procedure of Bonferroni-Hochberg) were selected as overrepresented. An analysis of cellular processes influenced by the protein profile obtained was performed using PathwayStudio (v5.00) software (Ariadne Genomics, Inc., Rockville, MD). PathwayStudio includes an automated text-mining tool which enables the software to generate pathways from the PubMed database and other public sources. Each identified cellular process was confirmed through the PubMed/Medline hyperlink embedded in each node. The domain annotations were assigned using the Pfam database.

Authors' contributions

PZ performed the LC-MS/MS experiments, data analysis, as well as data interpretation and contributed to the writing of the manuscript. NJ and YG contributed to the revision of the manuscript. XG contributed to the bioinformatic processing of the data. YW performed the sample preparation. ZZ, RH and JS supervised the experimental design and contributed to the data interpretation and manuscript evaluation. All authors read and approved the final manuscript.

Additional material

Additional file 1

Identified mouse oocyte proteins. Contains a list of all identified proteins from mouse MII oocytes in this study. Peptide information and spectrum information are also included.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-348-S1.xls>]

Additional file 2

Data from single peptide-based identification of oocyte proteins. Contains MS/MS spectra and fragment tables of single peptide-based identifications.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-348-S2.pdf>]

Additional file 3

Data from single peptide-based identification of oocyte proteins. Contains MS/MS spectra and fragment tables of single peptide-based identifications.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-348-S3.pdf>]

Additional file 4

Data from single peptide-based identification of oocyte proteins. Contains MS/MS spectra and fragment tables of single peptide-based identifications.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-348-S4.pdf>]

Additional file 5

Proteins of particular interest. Contains a list of high-abundance proteins, proteins with high mRNA expression levels, and proteins common to both the oocyte and the ES cell proteome.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-348-S5.xls>]

Acknowledgements

This study was supported by grants from 973 Program (No. 2006CB701503), NSFC (No. 90919014) and Program for Changjiang Scholars and Innovative Research Team in University (No. IRT0631).

References

1. Soyol SM, Amleh A, Dean J: **FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation.** *Development* 2000, **127(21)**:4645-4654.
2. Dong J, Albertini DF, Nishimori K, Kumar TR, Lu N, Matzuk MM: **Growth differentiation factor-9 is required during early ovarian folliculogenesis.** *Nature* 1996, **383(6600)**:531-535.
3. Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery GW, et al.: **Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner.** *Nat Genet* 2000, **25(3)**:279-283.
4. Minami N, Suzuki T, Tsukamoto S: **Zygotic gene activation and maternal factors in mammals.** *J Reprod Dev* 2007, **53(4)**:707-715.
5. Morisato D, Anderson KV: **Signaling pathways that establish the dorsal-ventral pattern of the Drosophila embryo.** *Annu Rev Genet* 1995, **29**:371-399.
6. Newport J, Kirschner M: **A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription.** *Cell* 1982, **30(3)**:687-696.
7. Tong ZB, Gold L, Pfeifer KE, Dorward H, Lee E, Bondy CA, Dean J, Nelson LM: **Mater, a maternal effect gene required for early embryonic development in mice.** *Nat Genet* 2000, **26(3)**:267-268.
8. Burns KH, Viveiros MM, Ren Y, Wang P, DeMayo FJ, Frail DE, Eppig JJ, Matzuk MM: **Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos.** *Science* 2003, **300(5619)**:633-636.
9. Payer B, Saitou M, Barton SC, Thresher R, Dixon JP, Zahn D, Colledge WH, Carlton MB, Nakano T, Surani MA: **Stella is a maternal effect gene required for normal early development in mice.** *Curr Biol* 2003, **13(23)**:2110-2117.
10. Esposito G, Vitale AM, Leijten FP, Strik AM, Koonen-Reemst AM, Yurttas P, Robben TJ, Coonrod S, Gossen JA: **Peptidylarginine deiminase (PAD) 6 is essential for oocyte cytoskeletal sheet formation and female fertility.** *Mol Cell Endocrinol* 2007, **273(12)**:25-31.
11. Li L, Baibakov B, Dean J: **A subcortical maternal complex essential for preimplantation mouse embryogenesis.** *Dev Cell* 2008, **15(3)**:416-425.
12. Memili E, Peddinti D, Shack LA, Nanduri B, McCarthy F, Sagirkaya H, Burgess SC: **Bovine germinal vesicle oocyte and cumulus cell proteomics.** *Reproduction* 2007, **133(6)**:1107-1120.
13. Susor A, Ellederova Z, Jelinkova L, Halada P, Kavan D, Kubelka M, Kovarova H: **Proteomic analysis of porcine oocytes during in**

- in vitro maturation reveals essential role for the ubiquitin C-terminal hydrolase-L1.** *Reproduction* 2007, **134(4)**:559-568.
14. Ellederova Z, Halada P, Man P, Kubelka M, Motlik J, Kovarova H: **Protein patterns of pig oocytes during in vitro maturation.** *Biol Reprod* 2004, **71(5)**:1533-1539.
 15. Calvert ME, Digilio LC, Herr JC, Coonrod SA: **Oolemmal proteomics identification of highly abundant heat shock proteins and molecular chaperones in the mature mouse egg and their localization on the plasma membrane.** *Reprod Biol Endocrinol* 2003, **1**:27.
 16. Vitale AM, Calvert ME, Mallavarapu M, Yurttas P, Perlin J, Herr J, Coonrod S: **Proteomic profiling of murine oocyte maturation.** *Mol Reprod Dev* 2007, **74(5)**:608-616.
 17. Ma M, Guo X, Wang F, Zhao C, Liu Z, Shi Z, Wang Y, Zhang P, Zhang K, Wang N, et al.: **Protein expression profile of the mouse metaphase-II oocyte.** *J Proteome Res* 2008, **7(11)**:4821-4830.
 18. Mann M, Jensen ON: **Proteomic analysis of post-translational modifications.** *Nat Biotechnol* 2003, **21(3)**:255-261.
 19. Marko-Varga G, Fehniger TE: **Proteomics and disease: the challenges for technology and discovery.** *J Proteome Res* 2004, **3(2)**:167-178.
 20. Lubec G, Krapfenbauer K, Fountoulakis M: **Proteomics in brain research: potentials and limitations.** *Prog Neurobiol* 2003, **69(3)**:193-211.
 21. Park YM, Kim JY, Kwon KH, Lee SK, Kim YH, Kim SY, Park GW, Lee JH, Lee B, Yoo JS: **Profiling human brain proteome by multi-dimensional separations coupled with MS.** *Proteomics* 2006, **6(18)**:4978-4986.
 22. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M: **The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins.** *Genome Biol* 2006, **7(9)**:R80.
 23. Schirle M, Heurtier MA, Kuster B: **Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry.** *Mol Cell Proteomics* 2003, **2(12)**:1297-1305.
 24. Wassarman PM, Jovine L, Litscher ES: **Mouse zona pellucida genes and glycoproteins.** *Cytogenet Genome Res* 2004, **105(24)**:228-234.
 25. Eppig JJ, Wigglesworth K, Pendola FL: **The mammalian oocyte orchestrates the rate of ovarian follicular development.** *Proc Natl Acad Sci USA* 2002, **99(5)**:2890-2894.
 26. Keller A, Nesvizhskii AI, Kolker E, Aebersold R: **Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search.** *Anal Chem* 2002, **74(20)**:5383-5392.
 27. Petyuk VA, Qian WJ, Hinault C, Gritsenko MA, Singhal M, Monroe ME, Camp DG 2nd, Kulkarni RN, Smith RD: **Characterization of the mouse pancreatic islet proteome and comparative analysis with other mouse tissues.** *J Proteome Res* 2008, **7(8)**:3114-3126.
 28. Al-Shahrour F, Carbonell J, Minguéz P, Goetz S, Conesa A, Tarraga J, Medina I, Alloza E, Montaner D, Dopazo J: **Babelomics: advanced functional profiling of transcriptomics, proteomics and genomics experiments.** *Nucleic Acids Res* 2008:W341-346.
 29. De La Fuente R: **Chromatin modifications in the germinal vesicle (GV) of mammalian oocytes.** *Dev Biol* 2006, **292(1)**:1-12.
 30. Cho CK, Shan SJ, Winsor EJ, Diamandis EP: **Proteomics analysis of human amniotic fluid.** *Mol Cell Proteomics* 2007, **6(8)**:1406-1415.
 31. Van Hoof D, Passier R, Ward-Van Oostwaard D, Pinkse MW, Heck AJ, Mummery CL, Krijgsvelde J: **A quest for human and mouse embryonic stem cell-specific proteins.** *Mol Cell Proteomics* 2006, **5(7)**:1261-1273.
 32. Ko YG, Nishino K, Hattori N, Arai Y, Tanaka S, Shiota K: **Stage-by-stage change in DNA methylation status of Dnmt1 locus during mouse early development.** *J Biol Chem* 2005, **280(10)**:9627-9634.
 33. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, et al.: **Large-scale analysis of the human and mouse transcriptomes.** *Proc Natl Acad Sci USA* 2002, **99(7)**:4465-4470.
 34. Nakayama KI, Nakayama K: **Ubiquitin ligases: cell-cycle control and cancer.** *Nat Rev Cancer* 2006, **6(5)**:369-381.
 35. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH: **Viable offspring derived from fetal and adult mammalian cells.** *Nature* 1997, **385(6619)**:810-813.
 36. Cowan CA, Atienza J, Melton DA, Eggan K: **Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells.** *Science* 2005, **309(5739)**:1369-1373.
 37. Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T: **Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells.** *Curr Biol* 2001, **11(19)**:1553-1558.
 38. Latham KE, Schultz RM: **Embryonic genome activation.** *Front Biosci* 2001, **6**:D748-759.
 39. Seydoux G, Braun RE: **Pathway to totipotency: lessons from germ cells.** *Cell* 2006, **127(5)**:891-904.
 40. Richter JD: **Translational control during early development.** *Bioessays* 1991, **13(4)**:179-183.
 41. Horikawa M, Kirkman NJ, Mayo KE, Mulders SM, Zhou J, Bondy CA, Hsu SY, King GJ, Adashi EY: **The mouse germ-cell-specific leucine-rich repeat protein NALP14: a member of the NACHT nucleoside triphosphatase family.** *Biol Reprod* 2005, **72(4)**:879-889.
 42. Hamatani T, Falco G, Carter MG, Akutsu H, Stagg CA, Sharov AA, Dudekula DB, VanBuren V, Ko MS: **Age-associated alteration of gene expression patterns in mouse oocytes.** *Hum Mol Genet* 2004, **13(19)**:2263-2278.
 43. Bichi R, Shinton SA, Martin ES, Koval A, Calin GA, Cesari R, Russo G, Hardy RR, Croce CM: **Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression.** *Proc Natl Acad Sci USA* 2002, **99(10)**:6955-6960.
 44. Narducci MG, Fiorenza MT, Kang SM, Bevilacqua A, Di Giacomo M, Remotti D, Picchio MC, Fidanza V, Cooper MD, Croce CM, et al.: **TCL1 participates in early embryonic development and is overexpressed in human seminomas.** *Proc Natl Acad Sci USA* 2002, **99(18)**:11712-11717.
 45. DeRenzo C, Seydoux G: **A clean start: degradation of maternal proteins at the oocyte-to-embryo transition.** *Trends Cell Biol* 2004, **14(8)**:420-426.
 46. Takahashi K, Yamanaka S: **Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.** *Cell* 2006, **126(4)**:663-676.
 47. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R: **In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state.** *Nature* 2007, **448(7151)**:318-324.
 48. Tani T, Shimada H, Kato Y, Tsunoda Y: **Bovine oocytes with the potential to reprogram somatic cell nuclei have a unique 23-kDa protein, phosphorylated transcriptionally controlled tumor protein (TCTP).** *Cloning Stem Cells* 2007, **9(2)**:267-280.
 49. Kozioł MJ, Garrett N, Gurdon JB: **Tpt1 activates transcription of oct4 and nanog in transplanted somatic nuclei.** *Curr Biol* 2007, **17(9)**:801-807.
 50. Shevchenko A, Wilm M, Vorm O, Mann M: **Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels.** *Anal Chem* 1996, **68(5)**:850-858.
 51. Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT, et al.: **Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling.** *Cell* 2006, **125(1)**:173-186.
 52. Foster LJ, de Hoog CL, Zhang Y, Xie X, Mootha VK, Mann M: **A mammalian organelle map by protein correlation profiling.** *Cell* 2006, **125(1)**:187-199.
 53. Shi R, Kumar C, Zougman A, Zhang Y, Podtelejnikov A, Cox J, Wisniewski JR, Mann M: **Analysis of the mouse liver proteome using advanced mass spectrometry.** *J Proteome Res* 2007, **6(8)**:2963-2972.
 54. Wang H, Qian WJ, Chin MH, Petyuk VA, Barry RC, Liu T, Gritsenko MA, Mottaz HM, Moore RJ, Camp II DG, et al.: **Characterization of the mouse brain proteome using global proteomic analysis complemented with cysteinyl-peptide enrichment.** *J Proteome Res* 2006, **5(2)**:361-369.
 55. Cox B, Kislinger T, Wigle DA, Kannan A, Brown K, Okubo T, Hogan B, Jurisica I, Frey B, Rossant J, et al.: **Integrated proteomic and transcriptomic profiling of mouse lung development and Nmyc target genes.** *Mol Syst Biol* 2007, **3**:109.
 56. Dieguez-Acuna FJ, Gerber SA, Kodama S, Elias JE, Beausoleil SA, Faustman D, Gygi SP: **Characterization of mouse spleen cells by subtractive proteomics.** *Mol Cell Proteomics* 2005, **4(10)**:1459-1470.

57. Yu LR, Conrads TP, Uo T, Kinoshita Y, Morrison RS, Lucas DA, Chan KC, Blonder J, Issaq HJ, Veenstra TD: **Global analysis of the cortical neuron proteome.** *Mol Cell Proteomics* 2004, **3(9)**:896-907.
58. Baker MA, Hetherington L, Reeves GM, Aitken RJ: **The mouse sperm proteome characterized via IPG strip prefractionation and LC-MS/MS identification.** *Proteomics* 2008, **8(8)**:1720-1730.
59. Maziarz M, Chung C, Drucker DJ, Emili A: **Integrating global proteomic and genomic expression profiles generated from islet alpha cells: opportunities and challenges to deriving reliable biological inferences.** *Mol Cell Proteomics* 2005, **4(4)**:458-474.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

