npg

Original Article

The proteomic analysis of human neonatal umbilical cord serum by mass spectrometry

Hong-juan SONG[1,3], Ping ZHANG[2], Xue-jiang GUO[2], Lian-ming LIAO[3], Zuo-min ZHOU[2], Jia-hao SHA[2], Yu-gui CUI[3], Hui JI[1,*], Jia-yin LIU[1,3,*]

[1]China Pharmaceutical University, Nanjing 210009, China; [2]Jiangsu Province Laboratory of Reproductive Medicine, Nanjing 210000, China; [3]Center of Clinical Reproductive Medicine, First Affiliated Hospital, Nanjing Medical University, Nanjing 210036, China

Aim: To investigate the proteome composition and function of human neonatal arterial umbilical cord.
Methods: Serum proteomic analyses were performed on samples from both males and females by using a combination of techniques: (1) removal of six high-abundance proteins, (2) tryptic digestion of low-abundance proteins, (3) separation of peptide mixtures by reverse-phase high-performance liquid chromatography (RP-HPLC), and (4) peptide identification using electrospray ionization tandem mass spectrometry (ESI-MS/MS).
Results: A total of 837 non-redundant proteins were identified, with 213 male-specific and 239 female-specific proteins. Among them, 319 proteins were identified by at least 2 distinct peptides. The subcellular localization, function, and pathway involvement for each of the identified proteins were analyzed. A comparison of this neonatal proteome to that of adult serum proteome revealed novel biomarkers, such as alpha-fetoprotein and periostin that were specific to newborn infants.
Conclusion: These data will contribute to a better understanding of the composition of umbilical cord serum and aid the discovery of novel biomarkers for the prenatal diagnosis of fetal abnormalities.

Keywords: mass spectrometry; reference specimen; serum proteome; umbilical cord serum

Introduction

Blood serum contains thousands of small molecules, including salts, lipids, proteins, and sugars. Qualitative and quantitative changes in the levels of serum proteins can be related to physiological or pathological states[1–4], and serum markers have long been used for the clinical diagnosis and therapeutic monitoring of various diseases[2]. Proteomics is being extensively used to profile adult serum proteins, with a special focus on the identification of disease-related serum markers to facilitate the early diagnosis and treatment of disease[2, 3, 5]. Serum samples that are used in research are usually prepared from venal blood[6]. An equivalent comparison of the neonate proteome with the adult serum proteome poses significant challenges, since blood from newborn infants is very difficult to acquire by venipuncture. By contrast, umbilical cord blood (UCB) is easily obtained, with only a small risk to the donor. The protein profile and the levels of markers in UCB are related to the physiological or pathological state of the neonate. Many

studies have detected novel biomarkers for fetal abnormalities from UCB[7, 8]. Low levels of α-2-HS glycoprotein/fetuin-A have been associated with intrauterine growth restriction (IUGR) cases[8]. These changes might be responsible for the impaired function of fetuin-A, which could lead to deficient fetal growth, especially with respect to osteogenesis, and/or to the development of complications that are frequently seen later in the lives of IUGR cases. To date, there has been no description of the protein profile of neonatal UCB by ESI-MS/MS. The focus of this study was to profile and analyze the proteome of UCB in order to investigate the neonate-specific composition and function of serum.

Materials and methods

Preparation of pooled male and pooled female serum specimens
All samples were provided by the Nanjing Maternity and Child Health Care Hospital and were prepared according to the BD protocol (the Human Proteome Organization Plasma Proteome Project pilot phase), with some modifications. This study was approved by the ethics committee of the Nanjing Maternity and Child Health Care Hospital, and UCB samples were collected after written informed consent was obtained

from the parents.

The inclusion criteria of the subjects were as follows: (1) pregnant women who were aged between 23 and 30 years, were non-smokers, had no history of addiction to alcohol, and weighed between 40 and 70 kg; (2) no medications taken in the 7 days prior to delivery; (3) parents who were healthy and also tested negative for human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), and syphilis; (4) infants who were born via vaginal delivery with clear amniotic fluid; and (5) newborn infants who were healthy and also had Apgar scores of 10 at both 1 and 5 min, weighed between 2.5 and 4.0 kg, and had no abnormalities.

Approximately 20 mL of UCB was directly aspirated from a single umbilical cord artery without anticoagulant. The whole blood was allowed to clot for 2.5 h at 4°C, and the clotted material was removed by centrifugation at 2600×*g* for 15 min. The resulting serum was then centrifuged at 12 000×*g* for 5 min at 4°C to remove any remaining cells. A protease inhibitor cocktail (Pierce #Prod#78415, Rockford, USA) was added to the serum to prevent protein degradation. The pooled male and pooled female samples were created by the combination of 2.0-mL serum samples from each male or female donor, respectively. The pooled samples were then stored at -20 °C until use.

Depletion of the highly abundant serum proteins

The removal of high-abundance proteins from UCB was performed for two main reasons: (1) to prevent interference with the measurement of low-abundance proteins during MS and (2) to reduce the possibility that the MS-based sequence analysis of serum-derived peptides is complicated by regions of high-sequence variability found in the abundant serum immunoglobulins. Six high-abundance proteins, namely albumin, transferrin, haptoglobin, α-1-antitrypsin, IgA, and IgG, were selected for removal from the UCB serum samples using a multiple affinity removal column system (Agilent Technologies, Palo Alto, USA). Briefly, the crude serum was thawed, diluted five-fold with buffer A, pH 7.4 (product No 5185–5987; Agilent Technologies), and filtered through 0.22 µm filters (Agilent Technologies) by centrifugation at 16 000×*g* at room temperature for 1.5 min. The diluted serum samples were injected onto a Multiple Affinity Removal System HPLC column (Agilent Technologies) in buffer A at a flow rate of 0.25 mL/min for 9 min. The bound proteins were then eluted in buffer B at a flow rate of 1.0 mL/min for 3.5 min. All chromatographic fractionations were performed at room temperature on an HP1100 HPLC system with the automated sample injector set at 47 °C. The unbound (low-abundance) and bound (high-abundance) proteins were collected in Eppendorf tubes and stored at -20 °C for further analysis.

Protein concentration

The fractions that were collected from 6 males and 6 females were pooled, respectively, into two Microcon spin concentrators with a 5-kDa molecular weight cut-off (Millipore). The fractions were spun at 12 000×*g* at 47 °C for 2 h. Protein concentrations were estimated with a Bradford protein assay using bovine serum albumin (BSA) (BioRad, Hercules, CA, USA) as a protein standard. Each protein mixture was lyophilized for digestion.

One-dimensional SDS-PAGE and in-gel digestion

The extracted protein (100 mg) was dissolved in SDS-PAGE loading buffer, boiled for 5 min, and loaded onto a single lane of a 1-mm thick 10% polyacrylamide gel. After separation, the proteins were visualized by silver staining of the gel according to the published procedure[9], with the minor modification of using a sensitizing solution that lacked glutardialdehyde. Follow-up analysis by Coomassie brilliant blue stain indicated that these six proteins were almost entirely removed (Figure 1).
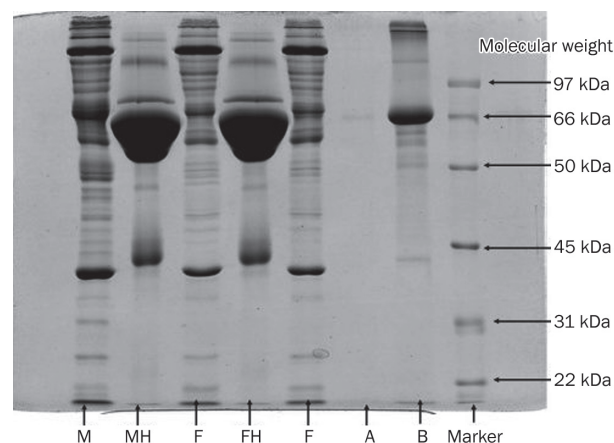

Figure 1. One-dimensional SDS-PAGE gel. F=low-abundance protein in females; FH=highly abundant protein in females; M=low-abundance protein in males; MH=highly abundant protein in males. Both A and B=other independent proteins.

The gel was then cut into 38 slices, which were subsequently cut into 1-mm³ gel particles and placed into 48 tubes for in-gel digestion. Briefly, the gel pieces were washed and destained in deionized water and then dehydrated. The gels were sequentially washed with 25 mmol/L ammonium bicarbonate and 50% acetonitrile (ACN) solution, followed by dehydration with 100% ACN and rehydration with 10 ng/µL trypsin (Promega, Madison, WI, USA) in 25 mmol/L ammonium bicarbonate. The gel pieces were incubated for 12 h at 37°C for protein digestion. The 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. The collected peptides were used for mass spectrometric analysis.

Online reversed-phase LC-MS/MS

For the capillary reversed-phase LC (cLC) and mass spectrometric (MS) analyses, the 48 fractions were sequentially loaded

onto a Michrom peptide CapTrap (M_W 0.5–50 kDa, 0.5×2 mm; Michrom BioResources, Auburn, USA) at a flow rate of 60 µL/min with buffer A (see below). The trap column effluent was then transferred to a reversed-phase microcapillary column (0.1×150 mm, packed with Magic C18, 5 µm, 100 Å; Michrom Bioresources, Auburn, USA). The reversed-phase separation of peptides was performed using the following buffers: 5% ACN, 0.1% formic acid in 95% ACN (buffer A), and 0.1% formic acid (buffer B). The samples were run on a 73-min gradient, which consisted of a 5%–45% buffer B for 55 min, 90% buffer B for 5 min, and 5% buffer B for 13 min. The peptide analysis was performed using the Finnigan LTQ ORBitrap (ThermoFinnigan, San Jose, USA) coupled directly to an LC column. The MS survey scan was obtained for the *m/z* range 400–1800, and the MS/MS spectra were acquired from the survey scan for the 10 most intense ions as determined by the 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.

Settings for peak list file generation and database searches

SEQUEST analysis software (Bioworks version 3.3; ThermoFinnigan) was used for the identification of peptide sequences[10]. DTA files (Bioworks version 3.3) for each MS/MS spectrum, with a minimum ion count of 8, were generated from the raw data for the peptide mass range of 400–8000. The International Protein Index (IPI) human database (ipi.HUMAN.v3.31; downloaded from ftp.ebi.ac.uk/pub/databases/IPI), which contains 67,511 entries was used in the SEQUEST analysis. The SEQUEST parameters were set as follows: peptide (parent ion) tolerance of 10 ppm, fragment ion tolerance of 1 Da, missed cleavages of 2 allowed, fixed modification of carbamidomethylation on Cys (+57 Da), and a differential modification of oxidation on Met (+16 Da). The criteria used for filtering peptides with low confidence scores were as follows: cross-correlation values (Xcorr) greater than 2.1 and 2.5 for doubly-charged ions and triply or higher charged ions, respectively; ΔCn values, representing the difference in Xcorr from the next highest value, of less than 0.1 were removed from the matched sequences; removal of the infrequent singly-charged ions. The false-positive error rate was calculated by analyzing all files with the same method, except against a "sequence-reversed" IPI Human database. The false-positive rate (FPR) was calculated as follows: FPR=# of false peptides/(# of true peptides+# of false peptides). The confidence score of the protein was calculated by applying the Peptide Prophet algorithm using the Scaffold software (v01_07_00; Proteome Software)[11].

Bioinformatics analysis

Functional annotations were performed using the 2008 DAVID software program (http://david.abcc.ncifcrf.gov/). DAVID is a web-based and client/server application that allows users to access a relational database of functional annotations for proteins. Entrez Gene IDs were entered into DAVID to determine whether there was any enrichment of gene ontology and Kyoto

Encyclopedia of Genes and Genomes (KEGG) pathways[12].

Results

Identification of neonatal umbilical cord serum proteins

The resulting peptides that were extracted from each gel slice were analyzed by automated reversed phase LC (RP-LC) coupled with MS/MS. From 48 LC-MS/MS runs, a total of 104,234 MS/MS spectra were acquired and searched against the IPI Human database v3.31. The resulting lists of proteins from each run were combined into a collective list of proteins with the Scaffold program. To experimentally test the FPR in our dataset, all output files were searched against the reversed sequence database. A peptide FPR of 5% and an 80% protein identification probability[13] were used as the cut-offs for Scaffold in order to generate a statistically valid list of proteins.

Since multiple unique peptide sequences can include identical protein sequences, determining the identities of proteins based upon sequenced peptides can occasionally be difficult unless specific peptides are identified. Therefore, shared information among the identified peptides was further analyzed. Proteins with shared peptides were organized into a single group, which was referred to as a protein group, in Scaffold. If a protein group comprised only isoforms or overlapping database entries that were indistinguishable by MS/MS analysis, then all of the proteins in this group were counted as a single protein. If these proteins were the products of distinct genes, then all the proteins in the group were discarded. As a result, the final number of identified proteins was lower than the actual value. These analyses resulted in the identification of 837 proteins that corresponded to 815 distinct genes.

Gene ontology (GO) analysis of the identified proteins

The proteins that were identified in this study were examined by GO analysis and categorized according to their cellular localization. The most abundant proteins were derived from the cytoplasm (29%), followed by membrane proteins (21%), extracellular region proteins (20%), unknown or unclassified proteins (15%), and nuclear proteins (12%) (Figure 2). In addition to cellular localization, the UCB proteins were also analyzed at the genome level (Figure 3). In addition, the UCB was enriched with proteins that were involved in cell communication pathways as well as in the complement and coagulation cascades (Figure 4)[14].

Comparison of male and female umbilical cord sera

The number of unique peptides is generally accepted as a semiquantitative measure of protein abundance[15, 16]. The proteins in male and female UCB were sorted into separate databases that included the unique peptides that were identified in each. The ten proteins with the most abundant number of unique peptides in males and in females are listed in Table 1. Apolipoprotein B, complement C3, and alpha-2-macroglobulin had the highest number of unique peptides, which indicated that the earlier depletion of the six most abundant proteins was efficient. The two datasets were compared in order to investigate the similarities and differences between male and
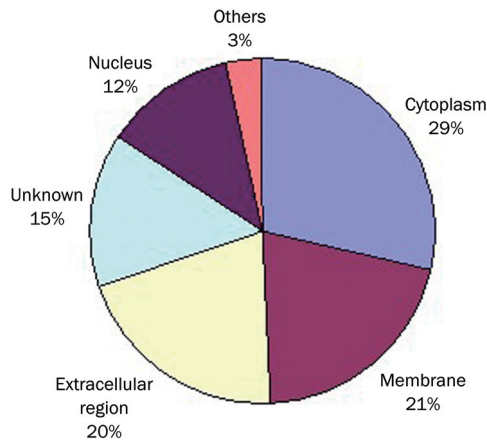
Figure 2. Classification of UCB proteins according to subcellular localization. The pie chart described the subcellular localizations of UCB proteins from the proteome. Subcellular localizations were assigned according to gene ontology annotations. When one protein was known to be localized in more than one cellular compartment, all of the localizations were counted non-exclusively.

Figure 3. UCB protein-associated genes. Proteins in UCB were found to be rich in many KEGG pathways. The numbers of proteins and significance of enrichment calculated by DAVID 2008 in each pathway were shown.

female UCB. Differentially expressed proteins were determined using the following criteria: (1) identification only in one dataset by at least two different peptides or (2) identification in one dataset by more than three times as many peptides in the other dataset[16]. There were 53 proteins that were differentially expressed between male and female UCB. The differential expression of these proteins could be a result of the distinct metabolic statuses between male and female fetuses.
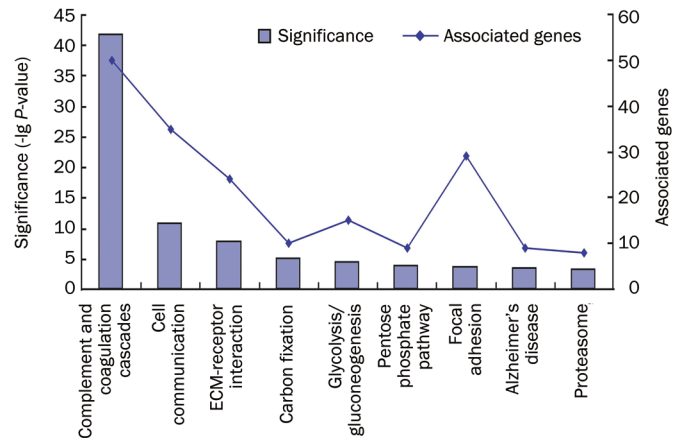
Comparison of neonatal umbilical cord serum and adult serum proteomes

To further investigate the highly expressed proteins in UCB and identify proteins that were specific to newborn infants, our protein dataset was compared with the adult plasma proteome. A total of 61 proteins, which were identified by two different peptides and were present in at least one gender, were present in the neonatal UCB serum proteome and not in the adult serum proteome (Table 2). Some of these proteins that are uniquely found in the UCB may be essential for fetal development.

Table 1. The ten most abundant proteins in UCB. The proteins from male and female UCB are sorted according to the unique peptides that were identified.

| Representative accession numbers	Protein name	Gender (Top 10)	Number of unique peptides (Female)	Number of unique peptides (Male)
IPI00478003	Alpha-2-macroglobulin precursor	Male, Female	75	69
IPI00745872	Isoform 1 of serum albumin precursor	Male, Female	48	52
IPI00022229	Apolipoprotein B-100 precursor	Male, Female	169	155
IPI00783987	Complement C3 precursor (Fragment)	Male, Female	98	83
IPI00032291	Complement C5 precursor	Male, Female	45	42
IPI00029739	Isoform 1 of Complement factor H precursor	Male, Female	45	45
IPI00017601	Ceruloplasmin precursor	Male, Female	39	39
IPI00019568	Prothrombin precursor (Fragment)	Male, Female	29	28
IPI00022418	Isoform 1 of Fibronectin precursor	Male, Female	57	49
IPI00292530	Inter-alpha-trypsin inhibitor heavy chain H1 precursor	Male, Female	32	27
IPI00022463	Serotransferrin precursor	Male, Female	39	39
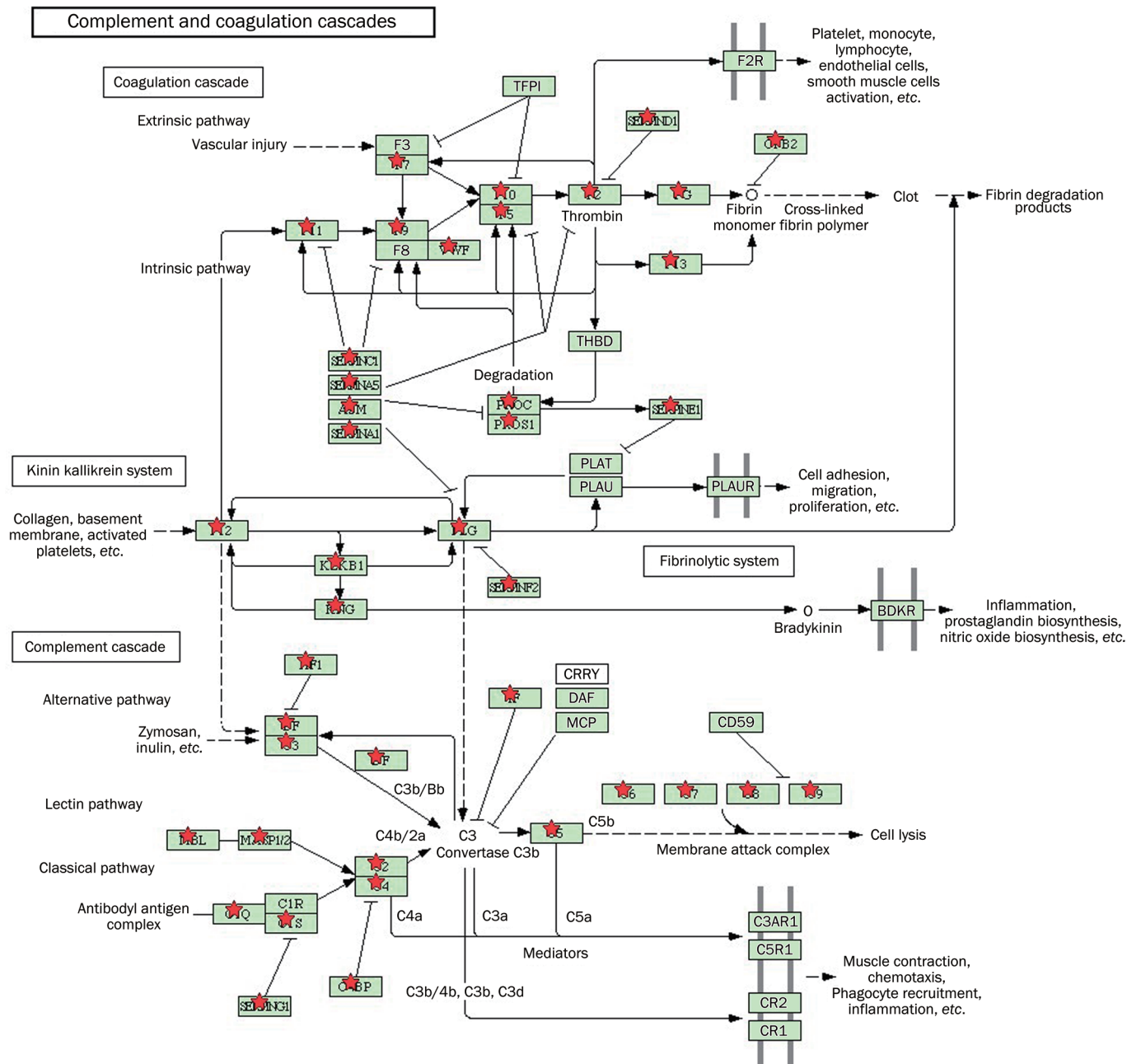
Figure 4. UCB proteins belonging to the complement and coagulation cascade pathways. Many of the proteins detected in UCB participate in the complement and coagulation cascade pathways according to KEGG pathway database. The proteins identified in UCB was marked by asterisks.

Discussion

The UCB proteome contains many high-abundance proteins that have various housekeeping roles and a number of secreted or shed low-abundance proteins that are critical for signaling cascades and regulatory events. During fetal development, regulatory factors can influence the fetus via the UCB. Meanwhile, metabolic products from the fetus may be released into UCB. Given that these components are present in UCB, a proteomic approach to identify biomarkers for these processes is highly merited. This study revealed that there were differences between the proteomic profiles of male and female fetuses and between those of the fetal and adult serums.

The proteins that belong to the UCB serum proteome are produced and secreted by either the fetus or the placenta. Many of the proteins that are present in UCB are of great interest, since their expression levels can reflect certain physiological and pathological conditions of the fetus and/or the status of the pregnancy. One protein of interest is alpha-fetoprotein (AFP), since its elevation in maternal serum correlates with several fetal abnormalities, including spina bifida, Down's syndrome, trisomy 13, and trisomy 18[7]. Our proteomic analysis identified several putative or currently used biomarkers for

Table 2. The 61 neonate-specific proteins in UCB.   A total of 61 proteins were identified in the UCB that were not present in the adult plasma proteome. Each neonate-specific protein was identified by two different peptides in at least one gender.

Accession numbers	Protein name	Number of unique peptides (Female)	Number of unique peptides (Male)
IPI00410241	POSTN Periostin, osteoblast specific factor	16	
IPI00430808	IGKC IGKC protein	11	13
IPI00293665	KRT6B Keratin, type II cytoskeletal 6B	7	4
IPI00156171	ENPP2 Isoform 1 of Ectonucleotide pyrophosphatase/phosphodiesterase family member 2 precursor	6	3
IPI00031008			
IPI00220213	TNC Isoform 1 of Tenascin precursor	5	7
IPI00399007	IGHG2 Hypothetical protein DKFZp686I04196 (Fragment)	5	5
IPI00021263			
IPI00789337	YWHAZ 14-3-3 protein zeta/delta	5	3
IPI00303482	FGL1 Fibrinogen-like 1 precursor	4	4
IPI00299485	CD93 Complement component C1q receptor precursor	4	3
IPI00014592	CHAD Chondroadherin precursor	4	3
IPI00009276	PROCR Endothelial protein C receptor precursor	4	3
IPI00220301	PRDX6 Peroxiredoxin-6	4	2
IPI00550363			
IPI00644531			
IPI00647915	TAGLN2 Transgelin-2	4	2
IPI00386879	IGHA1 CDNA FLJ14473 fis, clone MAMMA1001080, highly similar to Homo sapiens SNC73 protein (SNC73) mRNA	4	2
IPI00215914			
IPI00215917	ARF1 ADP-ribosylation factor 1	4	
IPI00032190	IHH Indian hedgehog protein precursor	3	4
IPI00845354	IGKC IGKC protein	3	4
IPI00002732	EXTL2 EXTL2 protein (Fragment)	3	4
IPI00019038	LYZ Lysozyme C precursor	3	3
IPI00291922	PSMA5 Proteasome subunit alpha type 5	3	1
IPI00009851	PCDH12 PCDH12 protein	3	1
IPI00291005	MDH1 Malate dehydrogenase, cytoplasmic	3	1
IPI00020990	OMD Osteomodulin precursor	3	1
IPI00745745	FUCA1 fucosidase, alpha-L-1, tissue		
IPI00843910		3	1
IPI00430847	IGKC IGKC protein	2	3
IPI00291136	COL6A1 Collagen alpha-1(VI) chain precursor	2	3
IPI00011194	FGFBP2 Fibroblast growth factor-binding protein 2 precursor	2	2
IPI00643041			
IPI00792352			
IPI0079567	RAN GTP-binding nuclear protein		
IPI00796462	Ran	2	2
IPI00017672	NP CDNA FLJ25678 fis, clone TST04067, highly similar to PURINE NUCLEOSIDE PHOSPHORYLASE	2	2
IPI00784430	IGKV3D-11 Similar to Ig kappa chain V-III region VG precursor	2	2
IPI00328703	OAF Out at first protein homolog precursor	2	2
IPI00031708	FAH Fumarylacetoacetase	2	2
IPI00014424	EEF1A2 Elongation factor 1-alpha 2	2	2
IPI00170867	DAXX Isoform 1 of Death domain-associated protein 6	2	1
IPI00220885			
IPI00736885			
IPI00827773	LOC440786 Ig kappa chain V-II region TEW	2	1
IPI00003802	MAN2A1 Alpha-mannosidase 2	2	1

Accession numbers	Protein name	Number of unique peptides (Female)	Number of unique peptides (Male)
IPI00009950	LMAN2 Vesicular integral-membrane protein VIP36 precursor	2	1
IPI00215767	B4GALT1 Isoform Long of Beta-1,4-galactosyltransferase 1	2	
IPI00759755		1	
IPI00218493	HPRT1 Hypoxanthine-guanine phosphoribosyltransferase	2	1
IPI00306311			
IPI00742980	PLEK Pleckstrin	2	
IPI00030205	IGKV3-20 Ig kappa chain V-III region HAH precursor	2	
IPI00736041	SERPINA11 similar to Serpin A11 precursor	2	
IPI00028908	NID2 Nidogen-2 precursor	2	
IPI00442073	CSRP1 Cysteine and glycine-rich protein 1	2	
IPI00303476			
IPI00789157	ATP5B ATP synthase subunit beta, mitochondrial precursor	1	3
IPI00791052			
IPI00746963	IGKC IGKC protein	1	2
IPI00382606	F7 Factor VII active site mutant immunoconjugate	1	2
IPI00295414	COL15A1 Collagen alpha-1(XV) chain precursor	1	2
IPI00784142			
IPI00010896			
IPI00554767	CLIC1 Chloride intracellular channel protein 1	1	2
IPI00334715			
IPI00718985	GRLF1 Isoform 1 of Glucocorticoid receptor DNA-binding factor 1	1	2
IPI00007960	POSTN Isoform 1 of Periostin precursor		14
IPI00641231			
IPI00418413	FLJ32569 hypothetical protein LOC148811		4
IPI00186460			
IPI00383762			
IPI00395679	COL2A1 Isoform 1 of Collagen alpha-1(II) chain precursor		3
IPI00219525			
IPI00747533	PGD 6-phosphogluconate dehydrogenase, decarboxylating		2
IPI00550488			
IPI00744692	TALDO1 similar to transaldolase 1		2
IPI00019176	RARRES2 Retinoic acid receptor responder protein 2 precursor		2
IPI00219446	PEBP1 Phosphatidylethanolamine-binding protein 1		2
IPI00003815			
IPI00793767			
IPI00794402	ARHGDIA Rho GDP-dissociation inhibitor 1		2
IPI00796541			
IPI00472961	IGKC IGKC protein		2
IPI00019954			
IPI00788184	CST6 Cystatin-M precursor		2
IPI00410402	TUBA3E Tubulin alpha-3E chain		2

the status of the fetus: glyceraldehyde 3-phosphate dehydrogenase, collagen alpha 1 (I), collagen alpha 1 (III), and HSPG2, which are used to test for Down's syndrome[14]; AFP, which is used for trisomy 13 and trisomy 18; fibronectin and plasminogen activator inhibitor, which are used for pre-eclampsia; and fibronectin[15, 17–19] and apolipoprotein A-I, which are used for cell membrane rupture[7]. Our results strongly indicate that a quantitative shotgun proteomic analysis of UCB is a feasible and effective method to screen for multiple pathologies.

A major challenge and bottleneck in proteomics lies between protein identification and target validation. Proteins with potential as biomarkers need to undergo further comparative analysis using samples from unaffected and affected pregnancies. In addition, markers that are currently used individually often have insufficient detection rates and specificities, which therefore requires the use of multiple markers. As an example, the prenatal screen for Down's syndrome currently includes three to five different markers. An integrated test or a fingerprinting assay is currently thought to potentially reflect the health status of a fetus better than a test involving a single marker. The MS-derived, systematic expression profiling of proteins from the UCB may serve as one such effective screening tool.

Our analysis identified 53 proteins that are differently expressed between male and female UCB. Of these, 28 proteins were higher in the female UCB, while 25 were higher in the male. Haptoglobin (Hp), which is a hemoglobin(Hb)-binding plasma protein, was identified by seven unique peptides in female UCB but not at all in male UCB. Hp is an acute-phase protein that responds to interleukin 6-type cytokines and is synthesized in the liver and, to some extent, in fat tissue[19]. There are two major allelic variants of Hp, which are Hp1 and Hp2[20]. Hp1 has higher Hb binding[21] and antioxidant capacities than Hp2[22–24]. Individuals carrying the Hp1 allele exhibit a lower incidence of angiopathies[25]. Gestational diabetes mellitus (GDM), which is a type of diabetes that is usually confined to the time of gestation, can be an early manifestation of type II diabetes since it increases the risk of developing type II diabetes later in life. Therefore, the high expression level of Hp in female UCB is of great interest, and, although our analysis could not discriminate between Hp1 and Hp2, further efforts should be made to examine this observation.

Proteins that are identified in the UCB serum proteome that are also absent from the adult serum proteome may reflect the specific metabolic status of the fetus and correlate with potentially important regulatory factors for the fetus. Periostin is one of these neonatal specific proteins that has a high expression level in UCB. Periostin, which was originally named osteoblast-specific factor-2 (Osf2), was first identified in bone and was shown to regulate the adhesion and differentiation of osteoblasts[26, 27]. Recently, periostin has been reported to be frequently over-expressed in various human cancers, since it stimulates metastatic growth by promoting cancer cell survival, invasion, and angiogenesis[28]. The expression of periostin in UCB is consistent with the marked rate of growth and development of the fetus.

Author contribution

Hong-juan SONG, Ping ZHANG and Xue-jiang GUO designed and performed the research; Xue-jiang GUO, Jia-hao SHA, Zuo-min ZHOU analyzed the data; Hong-juan SONG, Ping ZHANG, Xue-jiang GUO, Lian-ming LIAO, Yu-gui CUI, Hui JI, and Jia-yin LIU wrote the manuscript.

References

1 Adkins JN, Varnum SM, Auberry KJ, Moore RJ, Angell NH, Smith RD, *et al*. Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. Mol Cell Proteomics 2002; 1: 947–55.

2 Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002; 1: 845–67.

3 Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, *et al*. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002; 359: 572–7.

4 Pieper R, Gatlin CL, Makusky AJ, Russo PS, Schatz CR, Miller SS, *et al*. The human serum proteome: display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. Proteomics 2003; 3: 1345–64.

5 Aivado M, Spentzos D, Germing U, Alterovitz G, Meng XY, Grall F, *et al*. Serum proteome profiling detects myelodysplastic syndromes and identifies CXC chemokine ligands 4 and 7 as markers for advanced disease. Proc Natl Acad Sci USA 2007; 104: 1307–12.

6 He P, He HZ, Dai J, Wang Y, Sheng QH, Zhou LP, *et al*. The human plasma proteome: analysis of Chinese serum using shotgun strategy. Proteomics 2005; 5: 3442–53.

7 Vuadens F, Benay C, Crettaz D, Gallot D, Sapin V, Schneider P, *et al*. Identification of biologic markers of the premature rupture of fetal membranes: proteomic approach. Proteomics 2003; 3: 1521–5.

8 Karamessinis PM, Malamitsi-Puchner A, Boutsikou T, Makridakis M, Vougas K, Fountoulakis M, *et al*. Marked defects in the expression and glycosylation of alpha2-HS glycoprotein/fetuin-A in plasma from neonates with intrauterine growth restriction: proteomics screening and potential clinical implications. Mol Cell Proteomics 2008; 7: 591–9.

9 Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 1996; 68: 850–8.

10 Yates JR 3rd, Eng JK, McCormack AL, Schieltz D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal Chem 1995; 67: 1426–36.

11 Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003; 75: 4646–58.

12 Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, *et al*. DAVID: database for annotation, visualization, and integrated

discovery. Genome Biol 2003; 4: P3.

13 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: 5383–92.

14 Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, *et al*. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008; 36: D480–4.

15 Tsangaris GT, Karamessinis P, Kolialexi A, Garbis SD, Antsaklis A, Mavrou A, *et al*. Proteomic analysis of amniotic fluid in pregnancies with Down syndrome. Proteomics 2006; 6: 4410–9.

16 Van Hoof D, Passier R, Ward-Van Oostwaard D, Pinkse MW, Heck AJ, Mummery CL, *et al*. A quest for human and mouse embryonic stem cell-specific proteins. Mol Cell Proteomics 2006; 5: 1261–73.

17 Stubbs TM, Lazarchick J, Horger EO 3rd. Plasma fibronectin levels in preeclampsia: a possible biochemical marker for vascular endothelial damage. Am J Obstet Gynecol 1984; 150: 885–7.

18 Gao M, Nakabayashi M, Sakura M, Takeda Y. The imbalance of plasminogen activators and inhibitor in preeclampsia. J Obstet Gynaecol Res 1996; 22: 9–16.

19 Heinrich PC, Castell JV, Andus T. Interleukin-6 and the acute phase response. Biochem J 1990; 265: 621–36.

20 Bowman BH, Kurosky A. Haptoglobin: the evolutionary product of duplication, unequal crossing over, and point mutation. Adv Hum Genet 1982; 12: 189–261, 453–4.

21 Okazaki T, Nagai T. Difference in hemoglobin-binding ability of polymers among haptoglobin phenotypes. Clin Chem 1997; 43:

2012–3.

22 Langlois MR, Delanghe JR, De Buyzere ML, Bernard DR, Ouyang J. Effect of haptoglobin on the metabolism of vitamin C. Am J Clin Nutr 1997; 66: 606–10.

23 Melamed-Frank M, Lache O, Enav BI, Szafranek T, Levy NS, Ricklis RM, *et al*. Structure-function analysis of the antioxidant properties of haptoglobin. Blood 2001; 98: 3693–8.

24 Bernard D, Christophe A, Delanghe J, Langlois M, De Buyzere M, Comhaire F. The effect of supplementation with an antioxidant pre-paration on LDL-oxidation is determined by haptoglobin polymorphism. Redox Rep 2003; 8: 41–6.

25 Delanghe J, Cambier B, Langlois M, De Buyzere M, Neels H, De Bacquer D, *et al*. Haptoglobin polymorphism, a genetic risk factor in coronary artery bypass surgery. Atherosclerosis 1997; 132: 215–9.

26 Horiuchi K, Amizuka N, Takeshita S, Takamatsu H, Katsuura M, Ozawa H, *et al*. Identification and characterization of a novel protein, periostin, with restricted expression to periosteum and periodontal ligament and increased expression by transforming growth factor beta. J Bone Miner Res 1999; 14: 1239–49.

27 Litvin J, Selim AH, Montgomery MO, Lehmann K, Rico MC, Devlin H, *et al*. Expression and function of periostin-isoforms in bone. J Cell Biochem 2004; 92: 1044–61.

28 Kudo Y, Siriwardena BS, Hatano H, Ogawa I, Takata T. Periostin: novel diagnostic and therapeutic target for cancer. Histol Histopathol 2007; 22: 1167–74.