Prediction of cardiovascular risk in preterm neonates through urinary proteomics: An exploratory study

Estela Cabral\textsuperscript{a}, Henrique Soares\textsuperscript{b}, Hercílio Guimarães\textsuperscript{b,c}, Rui Vitorino\textsuperscript{c,e,*}, Rita Ferreira\textsuperscript{d}, Tiago Henriques-Coelho\textsuperscript{b,c}

\textsuperscript{a} School of Health Sciences, University of Minho, Braga, Portugal
\textsuperscript{b} Departamento de Pediatria, Faculdade de Medicina, Universidade do Porto, Porto, Portugal
\textsuperscript{c} Departamento de Cirurgia e Fisiologia, Faculdade de Medicina, Universidade do Porto, Porto, Portugal
\textsuperscript{d} QOPNA, Department of Chemistry, University of Aveiro, Aveiro, Portugal
\textsuperscript{e} iBiMED, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal

\section*{Introduction}

Low birth weight is the leading cause of morbidity and mortality in the neonatal period and in childhood overall. Over recent decades both the incidence of preterm birth and survival rates of preterm infants has increased, with babies born as early as 25 weeks gestation now having near 80\% chance to survive. Due to the immaturity of the organs at the time of birth, preterm infants exhibit an increased risk of developing a number of postnatal complications, namely cardiovascular and kidney diseases.\textsuperscript{1}

Cardiovascular diseases (CVD) are the leading cause of mortality in developed countries, with over 4 million deaths each year.\textsuperscript{2} For many years it was thought that the cardiovascular condition was determined solely by genetic factors and lifestyle, which includes amount of physical activity and quality of nutrition.\textsuperscript{3} In 1989, Barker et al.\textsuperscript{4} established a direct correlation between low birth weight and CVD in adulthood. This included hypertension and cardiovascular mortality. In the last decades, this discovery was confirmed and deeply explored with research being conducted on both animals and humans. The data collected suggests that the complex interaction between genetic constitution and the prenatal and early postnatal environment determines the growth and development of the fetus and defines the susceptibility to certain disorders in adult life, like hypertension, diabetes, dyslipidemia and coagulation disorders.\textsuperscript{5,6} Other studies have shown a relationship between preterm birth and cardiovascular diseases, namely, increased risk of cerebrovascular disease in young adulthood.\textsuperscript{7}

\begin{thebibliography}{9}
\end{thebibliography}
elevated blood pressure in adolescent girls and an atherogenic lipid profile in boys, metabolic syndrome in later life, increased plasma low-density lipoprotein, altered lipid and apolipoprotein metabolism with increased atherogenic indices at birth, increased risk of atrial fibrillation, and later-life risk of type 2 diabetes.

The development of proteomic techniques based on mass spectrometry has prompted the search for novel protein markers for many diseases. Proteomics allow a high-throughput analysis of samples as biofluids with the identification and quantification of thousands of proteins and peptides modulated by a given pathophysiological condition that might be seen as potential biomarkers. Urine is one of the most interesting and useful biofluids for routine testing and provides an excellent resource for the identification of novel biomarkers, with the advantage over blood samples due to the ease and less invasive nature of collection. Urine proteomics have been used to identify biomarkers associated with several diseases in neonates, such as bronchopulmonary dysplasia and necrotizing enterocolitis. For the time being, no specific urine biomarkers have been identified that may correlate with the increased risk of CVD in preterm infants.

The goal of the present study was to better understand the biological mechanisms by which prematurity is associated with increased risk for CVD, and identify potential biomarkers that would facilitate the development of effective screening and therapeutic strategies.

Methods

Characterization of the study population

Eight neonates followed by the neonatology service of Hospital São João – Porto were enrolled in the present study. The protocol was approved by the local ethics committee and the parents of all participating neonates gave their informed consent.

Neonates from the PreTerm group (n = 4) presented at birth a body weight of 1.23 ± 0.14 kg whereas all neonates from the Term group (n = 4) showed a body weight higher than 3.00 kg. Mean gestational age in the PreTerm group was 29 ± 1 weeks. All neonates of the Term group had a gestational age ≥37 weeks. Female and male neonates were included in both groups. Neonates from PreTerm group were given multivitamin and iron supplements. PreTerm neonates were also under parenteral nutrition, were submitted to phototherapy for jaundice, and received antibiotics for infectious risk. These neonates represent the typical clinical profile of PreTerm infants.

Sample collection and urine examination

Urine was collected in the first 20 days after birth by a non-invasive approach using sterile bags. Urine samples (approximately 10 ml) were then submitted to centrifugation for 2 min at 4000 × g at 4°C. In order to concentrate samples and remove salts, the supernatant was passed through 10-kDa filters (Vivaspin 500–10 kDa, Sartorius Biotech). The retentate was reconstituted in 100 µl of a solution of 6 M urea, 2 M thiourea, 1% CHAPS and 0.5% SDS. Total protein content was estimated using the RC-DC assay kit (BioRad). SDS-PAGE and in-gel digestion

Protein samples (40 µg) were separated by SDS-PAGE, using 12.5% polyacrylamide gels prepared as previously described. Gels were stained with BlueSafe® solution and were scanned with Gel Doc XR System (Bio-Rad). Samples were analyzed in duplicate. Following SDS-PAGE, complete lanes were cut out of the gel and sliced into 16 sections. Each section was in-gel digested with trypsin (0.01 µg/µL trypsin prepared in ammonium hydrogenocarbonate). Digestion process was quenched by the addition of 10% formic acid. Peptides were extracted from gel fragments by the addition of a solution of 10% formic acid and ultrapure acetonitrile (ACN; 1:1) and dried on a SpeedVac.

nanoLC–MS/MS analysis and protein ID

The dried extracted peptides were dissolved in 10 µL of mobile phase A (0.1% trifluoroacetic acid (TFA), 5% ACN, 95% water). All peptide mixtures were analyzed in two separate times. The tryptic digests were then separated using an Ultimate 3000 ( Dionex, Sunnyvale, CA) onto a 150 mm × 75 µm Pepmap100 C18 column with 3 µm particle size (Dionex, LC Packings) at a flow rate of 300 nL/min. The gradient started at 10 min and ramped to 50% Buffer B (85% ACN, 0.04% TFA) over a period of 45 min. The chromatographic separation was monitored at 214 nm using a UV detector (Dionex/LC Packings) equipped with a 3 nL flow cell. The peptides eluting from the column were mixed with a continuous flow of matrix solution (270 nL/min, 2 mg/mL α-Cyano-4-hydroxycinnamic acid in 70% ACN/0.3% TFA and internal standard Glu-Fib at 15 fmol) in a fraction microcolumn (Probot, Dionex/LC Packings) and directly deposited onto the LC-MALDI plates at 20 s intervals for each spot.

Samples were analyzed using a 4800 MALDI-TOF/TOF Analyzer (AbSciex), AS/N threshold of 50 was used to select peaks for MS/MS analyses. The spectra were processed using the TS2Mascot (v1.0, Matrix Science Ltd.) and submitted to Mascot software (v.2.1.0.4, Matrix Science Ltd.) for peptide/protein identification. Searches were performed in MASCOT using Homo sapiens release 13102014. Search was performed including data from all slices for global protein identification and emPAI calculation. A MS tolerance of 30 ppm was found for precursor ions and 0.3 Da for fragment ions, as well as two missed cleavages and methionine oxidation as variable modification. The confidence levels accepted for positive protein identification were p < 0.05. A false positive rate below 5% was obtained using a random database. Furthermore, proteins identified with one peptide were manually validated when MS/MS spectra presented at least 4 successive amino acids covered by b or y fragmentations.

Protein quantification and abundance measurement

The abundance of identified proteins was estimated by calculating the emPAI. The emPAI is an exponential form of PAI—(the number of detected peptides divided by the number of observable peptides per protein, normalized by the theoretical number of peptides expected via in silico digestion) defined as emPAI = 10PAI−1 and the corresponding protein content in mole percent is calculated as mol% = (emPAI/ΣemPAI) × 100. Microsoft Office Excel was used to calculate the mole percent. The theoretically observable peptides were determined by the in silico digestion of mature proteins using the output of the program Protein Digestion Simulator (http://panomics.pnnl.gov/software/). The observed peptides were unique parent ions including those with 2 missed cleavage. Mean protein emPAI values were log2 transformed for protein ratio calculation.

A CSV (comma separated values) dataset containing relevant information pertaining to all identified proteins was analyzed to extract meaningful information. An in-house developed C# program using Language Integrated Query (Microsoft Visual Studio 2012; Microsoft, USA) was used for data-mining the dataset. The output of the program has given several statistics which were then used by R Language scripts to produce emPAI distribution facilitating the dataset analysis.
Results

In order to characterize the urinary proteome profile, proteins from desalted urine samples were separated according to their molecular weight by SDS–PAGE. The large-scale LC–MS/MS analysis of all gel bands retrieved 434 unique proteins (Supplemental Table S1). According to PANTHER software, most of these proteins belong to the biological processes “metabolism”, “cellular processes” and “developmental processes”. Binding, catalytic activity and structural molecular activity are the most significant molecular functions of the identified proteins. Regarding cellular component, 30% of the identified proteins have extracellular origin.

In order to evaluate the effect of low birth weight on urinary protein profile, a label free quantification of GelLC–MS/MS data was performed. No differences were noticed in gel bands' profile among groups (Supplemental Figure S1). From the 236 proteins identified in the urine of at least two neonates, 126 were common to PreTerm and Term groups (Fig. 1), including, for instance, albumin, uromodulin and osteopontin. The analysis with PANTHER software demonstrated that most of the 58 proteins only identified in the urine of the Term group are related to catalytic and binding activities, structural molecular activity and receptor activity. Regarding the PreTerm group, most of the 37 unique proteins are involved in binding and catalytic activities, being receptor and enzyme regulator activities other important molecular functions.

Protein-protein interaction analysis with ClueGo and CluePedia, considering unique proteins identified in the PreTerm and Term groups was performed (Supplemental Figure S2). “Immune response”, “negative regulation of vasculature development” and “lipid storage and metabolism” are the most representative clusters in the PreTerm group. “Carbohydrate catabolic process”, “peptide metabolic process”, “forebrain neuron development” and “extracellular matrix organization” are the most representative clusters in the Term group.

Comparative abundance analysis evidenced alpha-1-antitrypsin, beta-2-microglobulin, retinol-binding protein 4, plasma protease C1 inhibitor, antithrombin-III and angiotensinogen as the most abundant proteins in the PreTerm group and cubulin, glutathione S-transferase P, calbindin, Ig kappa chain V and alpha-1-acid glycoprotein 1 as the most abundant ones in Term group (Fig. 2).

Protein–protein interaction analysis with ClueGo and CluePedia was performed, considering common proteins present in significant distinct levels (based on emPAI values) in the Term and PreTerm groups (Supplemental Figure S3). Major clusters of interacting proteins include those involved in “complement activation”, “blood coagulation”, “negative regulation of endopeptidase activity”, “retinol binding” and “heart growth”, which were up-regulated in PreTerm neonates. The clusters “negative regulation of apoptosis in response to DNA damage”, “regulation of chemokine production” and “kidney morphogenesis” were found down-regulated in this group of neonates.

Protein–protein interaction analysis performed with ClueGo and CluePedia considering unique proteins, and the ones in significant distinct levels highlights the up-regulation of “immune response” and “inflammation” in preterm neonates (Fig. 3).

Discussion

Our study highlights the impact of low birth weight–related clinical profile on urinary proteome, allowing a better understanding of the molecular pathways modulated by it, envisioning the identification of potential CVD biomarkers with diagnosis and prognosis significance.

Despite the low number of neonates enrolled in this exploratory study, data analysis highlight some biological processes modulated by prematurity, which validation with other methodologies in a higher cohort of neonates should be considered in the future. Metabolism was one of the biological processes altered in preterm neonates. Indeed, in PreTerm group lipid metabolism seems to prevail with proteins such as fatty acid-binding protein (FABP1), ribonuclease pancreatic (RNASE1), protein complement C3 and beta-hexosaminidase subunit beta (HEXB). It is known that newborns have minimum energy requirements for basal metabolism and growth that influence energy expenditure. Both energy and protein are necessary for a normal growth, and carbohydrates are the main source of energy for the brain and heart. Several weeks after birth, lipid oxidation develops. Most preterm newborns have greater needs of energy in order to approximate normal intrauterine growth and energy is mainly obtained from lipid metabolism. However, preterm infants have immature lipoprotein lipase activity (secondary to limited adipose tissue mass), which puts them at higher risk for hypertriglyceridemia. This metabolic adaptation observed might be a consequence of the nutritional support usually given at Neonate Intensive Care Unit in the early days of life for the necessary “catch-up” growth. However, rapid gains in adiposity have the potential to produce later life obesity, an increasingly recognized risk of excessive energy intake.

Upon birth, infants are exposed to relatively high concentrations of oxygen (O2) compared with intrauterine life. Under physiological conditions, blood oxygen saturation levels (Sao2) of the fetus average 45% to 55%. In the first few minutes after birth, blood oxygen content and O2 availability abruptly increase, which may alter oxygen-sensing pathways in the preterm neonates. The exposure to higher circulating oxygen concentrations elicits a burst of reactive oxygen species, which might be potentiated by iron supplementation. This usual therapeutic choice improves indicators of hematologic iron status and reduces the occurrence of iron-deficiency anemia in low birth weight infants; however, its effect on growth and neurodevelopment remains elusive. Some authors argue that the administration of ferrous iron might induce oxidative stress through the Fenton reaction, which might overload the immature antioxidant system of preterm neonates. Oxidative damage due to reactive oxygen species (ROS) generation has been linked to DNA damage and lipid peroxidation in premature newborns as well with the development of preterm associated diseases and adult diseases, including cardiovascular disease, cancer, chronic inflammation, and neurologic disorders.

Inflammation is intimately associated with increased oxidative stress which is present in preterm birth. Previous studies have shown increased inflammation markers in both mother
and newborn, including those born preterm.\textsuperscript{32} Oxidative stress and low-grade inflammation all contribute to the pathogenesis of chronic disorders, such as hypertension, cardiovascular and kidney diseases, and type 2 diabetes mellitus.\textsuperscript{25} In the preterm group, we observed an increase in lymphocyte mediated immunity, characterized by the unique proteins lysosome-associated membrane glycoprotein 1 (LAMP1) and C3, up-regulation of beta-2-microglobulin (B2MG) and hemopexin (HPX), and of proteins involved in complement activation (plasma protease C1 inhibitor (SERPING1) and mannan-binding lectin serine protease 2 (MASP2)). The activation of this pathway reflects the fact that preterm neonates are not fully immunocompetent at birth\textsuperscript{31} being more susceptible to infection than term neonates.\textsuperscript{34}

Altered cardiac growth, negative regulation of vasculature development, suppressed contractile function and enhanced responsiveness to stress are among the alterations reported in preterm neonates. These effects might contribute to the increased vulnerability of organ perfusion in preterm neonates.\textsuperscript{35} Indeed, a cluster that stands out in the PreTerm group is that involved in heart growth and cardiac muscle tissue growth. During maturation of the heart, fetal cardiomyocytes undergo both hyperplasia and hypertrophy. In humans, cardiomyocytes proliferate actively

**Fig. 2.** Normalized abundance of urinary proteins in PreTerm vs. Term groups. Protein accession number has correspondence to protein name at Supplemental Table S1.

**Fig. 3.** ClueGo and CluePedia analysis of protein-protein interaction considering unique proteins per group and common proteins present in significant distinct levels (based on emPAI values) in the urine of neonates from Term and PreTerm groups. Green nodes refer to biological processes positively regulated in preterm neonates whereas red nodes refer to the ones widespread in term neonates. As can be depicted, immune response, negative regulation of endopeptidase activity, regulation of defense response, ion homeostasis and catabolic processes are biological processes prevalent in preterm neonates.
until 36-week gestation after which maturation, differentiation, and growth by hypertrophy take place. Angiotensinogen (AGT) and retinol-binding protein 4 (RBP4) were found up-regulated in the PreTerm group, that are involved in heart growth. Increased levels of these proteins were already related to increased risk of cardiovascular disease. AGT is an essential component of the renin-angiotensin system, a potent regulator of blood pressure, body fluid and electrolyte homeostasis. Polymorphisms of AGT have been associated to an increased risk of heart diseases in newborns.

Although hypertension would not be caused by changes in cardiac structure and function after preterm birth per se, the heart of preterm-born neonates may be particularly susceptible to the effect of elevated blood pressure and associated risk factors for heart disease. The up-regulation of S100A9 further supports preterm-related increased risk of cardiovascular diseases, once was suggested to be an early marker for the detection of acute coronary syndromes and of cardiovascular risk in humans.

Taken together, in the present study we have provided new insights on the molecular pathways up-regulated in preterm neonates using urine proteomics, namely lymphocyte mediated immunity, metabolism adaptation, oxidative stress, negative regulation of vascular development and tissue remodeling. These pathways are related with cardiac development, supported by the presence of cardiac growth-related proteins in the PreTerm group, AGT and RBP4. Since immune response, oxidative stress and metabolic adaptation are intimately related with the development of cardiovascular diseases, the unique and upregulated proteins found in the PreTerm group may be seen as putative biomarkers to be included in multimarker strategies for the identification of neonates that are at higher risk of developing cardiovascular diseases. Future studies in a large cohort of preterm and term neonates as well as the follow-up of urinary proteome dynamics will be important to test the diagnostic value of these proteins.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by Portuguese Foundation for Science and Technology (FCT), European Union, QREN, FEDER and COMPETE for funding the UnIC, iBIMED and QOPNA research units (UID/IC/00051/2013, UID/BIM/04501/2013 and UID/QBI/00062/2013), RV (IF/00286/2015), and by RNEM (Portuguese Mass Spectrometry Network).

The authors thank to nurses’ team from Neonate Intensive Care Unit of Hospital São João for their support in collecting newborns’ urine samples.

Appendix A. Supplementary data


References

34. Moodley T, Vella C, Dhanabakhch O, Branford-White CJ, Arachidonic Craw- ford MA. docosahexaenoic acid deficits in preterm neonatal mononuclear


