Porto Biomed. J. 2017;xxx(xx):xxx–xxx

PORTO BIOMEDICAL JOURNAL
WHERE SCIENCE MEETS KNOWLEDGE

Porto Biomedical Journal

http://www.portobiomedicaljournal.com/

Original article

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

Estela Cabral[a], Henrique Soares[b], Hercília Guimarães[b,c], Rui Vitorino[c,e,*], Rita Ferreira[d], Tiago Henriques-Coelho[b,c]

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

ARTICLE INFO

*Article history:*
Received 6 March 2017
Accepted 10 April 2017
Available online xxx

*Keywords:*
Cardiovascular diseases
GeLC–MS/MS
Low birth weight
Neonates
Urine

ABSTRACT

Preterm birth has been associated with an increased risk of cardiovascular diseases (CVD) in adulthood. The goal of our study was to give new molecular insights on the relationship between prematurity and CVD risk and to identify putative biomarkers that would facilitate the development of effective screening and therapeutic strategies. In this sense, mass spectrometry (MS)-based proteomics was applied to the characterization of urine protein profile.

GeLC–MS/MS analysis of urine (desalted and concentrated with a 10-kDa filter) followed by bioinformatics was applied for the characterization of preterm and full-term neonates. Urine proteome profiling retrieved 434 unique proteins, from which 126 were common to both groups, 37 were unique to preterm and 58 to full-term neonates. Protein-protein interaction analysis for unique proteins and common ones present in significant distinct levels retrieved immune system, metabolism, defense systems and tissue remodeling as the most representative clusters in preterm neonates.

Metabolic adaptation along with the up-regulation of heart growth (identified by angiotensinogen and retinol-binding protein 4) may account for an increased CVD risk in preterm neonates. These proteins may have predictive value of CVD in adulthood of this specific group of neonates. The follow-up of urinary proteome dynamics of preterm and full-term neonates will be crucial for the validation of this hypothesis.



## 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.[1]

Cardiovascular diseases (CVD) are the leading cause of mortality in developed countries, with over 4 million deaths each year.[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.[3] In 1989, Barker et al.[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.[5,6] Other studies have shown a relationship between preterm birth and cardiovascular diseases, namely, increased risk of cerebrovascular disease in young adulthood,[7]

*Abbreviations:* AGT, angiotensinogen; CVD, cardiovascular diseases; GeLC, gel electrophoresis followed by liquid chromatography; MS, mass spectrometry; RBP4, retinol-binding protein 4.

* Corresponding author at: Department of Medical Sciences, Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.
*E-mail address:* rvitorino@ua.pt (R. Vitorino).

http://dx.doi.org/10.1016/j.pbj.2017.04.002

elevated blood pressure[8] in adolescent girls and an atherogenic lipid profile in boys,[9] metabolic syndrome in later life, increased plasma low-density lipoprotein,[10] altered lipid and apolipoprotein metabolism with increased atherogenic indices at birth,[11] increased risk of atrial fibrillation,[12] and later-life risk of type 2 diabetes.[13]

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.[14] 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.[15] Urine proteomics have been used to identify biomarkers associated with several diseases in neonates, such as broncopulmonary dysplasia[16] and necrotizing enterocolitis.[17] 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 $1.23 \pm 0.14$ 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 $29 \pm 1$ weeks. All neonates of the Term group had a gestational age $\geq$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 $4000 \times g$ 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 ressuspended in 100 μL of a solution of 6 M urea, 2 M tiourea, 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.[18] 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 microcollector (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). A S/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.[19] The emPAI is an exponential form of $PAI^{-1}$ (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 $\text{emPAI} = 10^{\text{PAI}} - 1$ and the corresponding protein content in mole percent is calculated as $\text{mol \%} = (\text{emPAI}/\Sigma\text{PemPAI}) \times 100$. 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.

Please cite this article in press as: Cabral E, et al. Prediction of cardiovascular risk in preterm neonates through urinary proteomics: An exploratory study. Porto Biomed. J. 2017. http://dx.doi.org/10.1016/j.pbj.2017.04.002