Journal of proteome research

Subscriber access provided by Imperial College London | Library

Technical Note

Development of a pipeline for exploratory metabolic profiling of infant urine

Frances Jackson, Nancy Georgakopoulou, Manuja R. Kaluarachchi, Michael Kyriakides, Nicholas James Andreas, Natalia Przysiezna, Matthew J Hyde, Neena Modi, Jeremy K. Nicholson, Anisha Wijeyesekera, and Elaine Holmes

J. Proteome Res., **Just Accepted Manuscript** • DOI: 10.1021/acs.jproteome.6b00234 • Publication Date (Web): 01 Aug 2016

Downloaded from http://pubs.acs.org on August 3, 2016

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

Journal of Proteome Research is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Development of a pipeline for exploratory metabolic profiling of infant urine

Frances Jackson[1], Nancy Georgakopoulou[1], Manuja Kaluarachchi[2], Michael Kyriakides[1], Nicholas Andreas[3], Natalia Przysiezna[3], Matthew J. Hyde[3], Neena Modi[3], Jeremy K. Nicholson[1,4], Anisha Wijeyesekera[1*] and Elaine Holmes[1,4]

[1]Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, South Kensington Campus, SW7 2AZ, United Kingdom

[2]Metabometrix Ltd, Bioincubator, Prince Consort Road, South Kensington, London, SW7 2AZ, United Kingdom

[3]Section of Neonatal Medicine, Department of Medicine, Imperial College London, Chelsea and Westminster Hospital Campus, SW10 9NH, United Kingdom

[4]MRC-NIHR National Phenome Centre, Department of Surgery and Cancer, Imperial College London, Hammersmith Hospital Campus, W12 0NN, United Kingdom

ABSTRACT

Numerous metabolic profiling pipelines have been developed to characterize the composition of human biofluids and tissues, the vast majority of these being for studies in adults. To accommodate limited sample volume and to take into account the compositional differences between adult and infant biofluids, we developed and optimized sample handling and analytical procedures for studying urine from newborns. A robust pipeline for metabolic profiling using NMR spectroscopy was established, encompassing sample collection, preparation, spectroscopic measurement and computational analysis. Longitudinal samples were collected from five infants from birth until 14 months of age. Methods of extraction, effects of freezing and sample dilution were assessed and urinary contaminants from breakdown of polymers in a range of diapers and cotton wool balls were identified and compared, including propylene glycol, acrylic acid and tert-butanol. Finally, assessment of urinary profiles obtained over the first few weeks of life revealed a dramatic change in

composition, with concentrations of phenols, amino acids and betaine altering systematically over the first few months of life. Therefore, neonatal samples require more stringent standardization of experimental design, sample handling and analysis compared to adult samples in order to accommodate the variability and limited sample volume.

Key words: Metabolic profiling, Metabonomics, NMR spectroscopy, Pipeline development, Infant, Urine

INTRODUCTION

The early life environment has been shown to exert a long-reaching impact on childhood and later life health. Mode of birth, feeding and drug exposure have all been associated with disease risk; including asthma and allergy in childhood, and metabolic syndrome and cardiovascular disease in later life[7].

Metabolic phenotyping of biofluids using high resolution spectroscopic techniques can be used to effectively characterize physiological or pathological processes with the aim of biomarker identification or augmenting mechanistic understanding of biology[1,2,3,4]. Collection and analysis of infant biofluids imposes an analytical challenge, since in contrast to adult samples, sample volumes are typically smaller. Urine samples are collected from neonates using plastic collection bags secured to the infant with adhesive, or more commonly, disposable diapers with sterile cotton balls or pads inside the diaper[18,19,20,] since the adhesive bags can cause skin irritation. Metabolic analysis of urine collected directly onto cotton wool or diapers can cause the introduction of contaminants from both the diaper and the cotton ball and the sample can also be contaminated by fecal material [21,22,23,24,25,34].

The clinical potential of metabonomics in neonatology is evidenced by a growing body of publications, demonstrating how physiological fetal development or pathological conditions affect perinatal and neonatal outcomes [7,8,9,10]. In addition to providing a detailed phenotype of acute infant illnesses, metabolic profiling can provide a baseline template for understanding downstream health. For this reason there is an increasing interest in establishing a phenotypic database of postnatal maturation and development of newborns. The literature relating to the biochemical composition of biofluids in neonates is relatively sparse with little known of the normal metabolic trajectory of the healthy neonate[11]. Growth, gestational age, nutritional composition of breast milk and infection have been shown to influence the metabolic composition of urine and plasma[12,13,14,15].

Thus far, metabonomic approaches have been used in experimental neonatology to shed light on conditions such as inborn errors of metabolism, cardiovascular, renal, and respiratory diseases; postnatal hypoxia/asphyxia, systemic infections (e.g. neonatal sepsis), brain injury, necrotizing enterocolitis, investigation of short- and long-term effects of prematurity, low birth weight and mode of delivery, as well as stratification of infants into risk groups according to response to various pharmacological treatments, such as gastroesophageal reflux treatments. However, sample collection and analysis methods have not been standardized and the studies to date have used different sample collection procedures, analytical platforms and data modelling strategies[5,6,7,9,11,13,14,15,16,17].

We present a systematic assessment of sample collection, processing and spectral analysis of urine samples from limited volume samples applied to newborn infants in order to characterize the changes in metabolic phenotype over the first few weeks of life, and to address the challenge of sample limitations and potential contamination from external sources. We propose an optimized analytical pipeline for NMR spectroscopic analysis of infant urine to ensure robust data generation and to facilitate inter-laboratory comparison of datasets.

MATERIALS

We developed and applied a sample collection and analysis pipeline for infant urine samples as defined in Figure 1. Urine collection was based around extraction of urine directly from cotton wool balls, or onto cotton wool pads inserted into diapers since clean catch urine, which is standard for adult collection, is difficult to obtain from infants. Contaminants deriving from the collection materials introduced were characterized in the urine metabolic profile. Urine samples were collected from infants born at term as part of an ongoing study at the Chelsea and Westminster Hospital. Research Ethics Committee approval was obtained from the West London Research Ethics Committee (REC reference number 12/LO/0203 and 10/H0/7135)

Extraction of urine from cotton balls:

A volume of 3 mL urine was added onto cotton wool and incubated for two hours at 37°C (to mimic the typical infant urine collection conditions). Urine was extracted from the cotton wool in the following ways; 1) urine filled cotton wool was frozen overnight at -80 °C prior to extracting urine; 2) extracted directly from the cotton wool immediately after collection and frozen overnight at -80°C; 3) extracted directly from the cotton wool immediately after collection; as well as 4) a fresh clean catch urine sample. Extraction from the cotton wool was carried out in two ways to assess whether chemical contamination originated from exposure to laboratory equipment: 1) squeezed from the cotton wool into a falcon tube or 2) extracted into a falcon tube by centrifugation. Extraction by centrifugation involved adding the cotton wool to a falcon tube and positioning an open ended plastic tube (a cryo-tube cut into a cylindrical shape, Supplementary Figure 1) underneath it to stop the cotton wool from moving to the bottom of the tube in order to ensure as much urine as possible was extracted. Diapers were also processed using a similar method. Urine was added to cotton wool and this was placed into the different diapers and incubated at 37°C for two hours after which urine was extracted by manually squeezing the cotton wool. For NMR analysis, 540µl of urine was transferred to 5mm NMR tubes, and 60µl of phosphate buffer added (Potassium Dihydrogen Phosphate (KH_2PO_4), 3-Trimethylsilyl Propionic Acid-d4 Acid Sodium Salt (TSP), Sodium Azide (NaN_3), and Potassium Hydroxide Solution KOH 45% w/w, Sigma-Aldrich, Gilliingham, UK). In addition to urine, water was used to obtain an experimental blank sample from both cotton wool and diapers in order to identify potential chemicals leaching from the collection materials.

Acquisition of ^{1}H NMR spectral profiles of urine and blank samples: All ^{1}H NMR spectra were acquired on a Bruker DRX-600 spectrometer (Bruker Biospin, Karlsruhe, Germany) operating at 600.29MHz for proton observation using a standard one dimensional water pre-saturation pulse sequence [relaxation delay-90°-t1-90°-tm-90°-acquire free induction decay (FID)]. The relaxation delay was 4 seconds; with application of a 90° radio frequency pulse, t1 (the interpulse delay), which

was set to 3µs, while tm is the mixing time of 100ms. The probe was matched and tuned automatically to the proton transmitter resonance frequency before acquisition for each sample. Samples were run at a temperature of 300K.

Spectral preprocessing and preliminary comparisons of the spectra were performed using the Bruker TopSpin 3.1 programme. Multivariate statistical modeling was performed using Mathworks Matlab R2014a (The Mathworks Inc., USA) and Umetrics SIMCAP 14.0 (Umetrics, Sweden).

Characterization of chemical contamination from collection materials:

To assess contamination of infant urine metabolic profile, different diapers and cotton wool brands were investigated. Three brands of diapers were assessed; for purpose of brevity these brands will be regarded as Diaper brand 1 (Pampers), Diaper brand 2 (UK hospital brand, Libero) and Diaper brand 3 (UK store brand, Tesco). Four cotton wool brands were also assessed; three UK store brands (Tesco, Sainsbury's and Boots) and UK hospital brand (Robinson's Healthcare). Since the capacity of urine and water as solvents may be different, the collection materials were assessed using H_2O as a blank sample as well as a pooled infant urine sample from infants at postpartum ages from 0-19 weeks. An aliquot of 3 mL of a quality control infant urine sample was transferred onto four different brands of cotton wool ball (Sainsbury's, Tesco, Boots and Robinson Healthcare) and also onto the three different diapers (Pampers, Tesco own brand, and Libero) combined with the cotton wool. Samples were analyzed using ^{1}H NMR spectroscopy.

Assessment of dilution factor and spectral normalization on infant urine profiles

Bio-samples collected from infants are often limited in amount compared with adults; this is especially true for preterm infants who, in addition to producing smaller volume of urine than equivalent term babies, often have other underlying conditions. Therefore, as part of method development for infant biofluid analysis, the detection limits in terms of obtaining reproducible urine profiles were investigated for metabolic phenotyping by ^{1}H NMR spectroscopy. Urine was

collected from five term-born infants in the first week postpartum and a dilution series (original concentration sample, and a dilution series of ½, ¼, ⅛ of the original concentration - as detailed in Supplementary Information - Figure S2) were created for each of the five samples. To compensate for the decrease in signal to noise in more dilute samples, the number of scans used to acquire spectral profiles was increased by √n, which in turn increased the duration of the NMR experiment. The signal to noise ratio (S/N) of an NMR signal is improved by using signaling averaging; the S/N improves as √n, where n is the number of signals averaged. NMR data were normalized using the probabilistic quotient normalization (PQN), which particularly accommodates for dilution differences between samples by scaling the spectra to a median reference spectrum[28.] Although the spectra of urine extracted from cotton wool or directly form diapers contained contaminants, the data were not remodeled after exclusion of these resonances since they do not dominate the coefficients of any model with the exception of the comparison of different cotton wool and diaper brands.

***Assessment of age variability on infant urine profiles*:**

Infants develop rapidly from birth with consequential adaptation of physiology, which is likely to cause their metabolic profile to dramatically change in the first few weeks of life. Thus, unlike adults where metabolic phenotypes are generally stable over narrow age ranges, the time of sampling may be more important in infant studies. Therefore, the assessment of age variability on infant profiles were investigated for metabolic phenotyping by ^{1}H NMR spectroscopy. Urine samples taken for mother-infant comparison were obtained from mother-infant dyads (n=10) using only infants born at term. Urine samples were taken from the infant at 0, 14 days, 8 weeks and 14 months post-partum and a further maternal sample was taken at the time of birth for comparison.

RESULTS AND DISCUSSION

Assessment of methods for extracting urine from cotton wool

The centrifugation apparatus designed (Supplementary information: Figure S1) allowed efficient extraction of urine from the cotton wool compared to manual squeezing. In contrast to the study by Goodpaster and colleagues[25], who reported no chemical leaching from cotton balls, we found clear evidence of tert-butanol, iso-butanol, acetone, formate and acetate leaching into both the water blank and urine sample blank which were extracted from the cotton. The urine sample extracted from the cotton wool after squeezing, had a much higher concentration of acetone, tert-butanol, and iso-butanol than the urine obtained via centrifugation, suggesting that manual squeezing of cotton wool results in greater destruction of the cotton wool fibers and therefore greater release of contaminants into the urine sample. This further underscores the need for well-defined standard operating procedures and materials and, although beyond the scope of the current study, it raises the question as to whether exposure of infants to these chemicals has any health-related consequences.

We further assessed the effect of freezing the urine sample - in the cotton wool prior to extraction as well as frozen after extraction - and compared this with obtaining urine fresh from a cotton wool ball and found that there were no obvious differences in the level of chemical contamination or endogenous metabolite profile. Therefore, we find freezing the urine containing cotton wool directly after collection at -80 °C (which is generally more practical than performing immediate extraction of the urine), to be an acceptable practice.

Furthermore, despite the introduction of contaminants from the cotton wool, the urine profiles were highly similar and a change in citrate concentration was the only visual difference in endogenous compound concentration after urine had been extracted from the cotton wool, which suggests that citrate may be partially retained in the cotton wool. Furthermore, the citrate signals (δ2.55 -CH_2; δ2.66 -CH_2) became broader after treatment with cotton wool which also suggests that there was binding of this metabolite (Figure 2(c)).

Characterization of chemical contamination from collection materials:

As expected, we found that sample collection, preparation and acquisition parameters had a strong effect on spectral composition and that the limitation of sample volume, constraints of sample collection and the temporal variation in biofluid composition in the newborn infant introduced greater chemical variation than would be typically found in adults. Therefore, it is essential to standardize methodological protocols as far as possible. Based on the results of this study we present some recommendations based on the experimental practice.

Clean catch urine is difficult to obtain from infants and thus typical collection means include placing adhesive sample collection bags on the infants, or more practically, relying on extracting urine directly from the diaper or from cotton wool balls placed inside the diaper. Both cotton wool and diapers contain chemical contaminants, which contribute to the urinary metabolic profile. NMR easily detects these contaminants within the urine sample. It is therefore necessary to recognize these contaminants so that their signals can be excluded from statistical modelling which aims to evaluate physiological or pathological processes. As expected, normalization to the median spectrum (Figure 3c) ameliorated the variation caused by differential dilution and contamination from collection material, but there is potential for a strong bias to be introduced if spectra are normalized to total intensity.

Visual comparison of the NMR spectra obtained from blank, original urine samples and cotton wool extracted urine samples were assessed for chemical contaminants (Figure 2a). From cotton wool, the major contaminants were identified as acetate (δ1.92,-CH_3), acetone (δ2.23,-CH_3), formate (δ 8.46,-CH), citrate (δ2.55 -CH_2; δ2.66 -CH_2) and propylene glycol (δ1.14 -CH_3; δ3.43 –CH; δ3.54 –CH; δ3.87 -CH_2), tert-butanol (δ1.25 -CH_3), iso-butanol (δ0.88 -CH_3) as well as a peak at δ6.02 (as yet unidentified).

Different cotton wool brands contained variable concentrations of these metabolites but the qualitative presence of the metabolites was consistent. Contaminants from diapers included propylene glycol (δ1.14 -CH_3; δ3.4–CH; δ3.6 –CH; δ3.9 -CH_2), N, N-dimethylacetamide (δ2.1 -CH_3, 2.9

-CH_3; 3.1 -CH_3) ethylene glycol (δ3.7 -CH), tert-butanol (δ1.25 -CH_3), isobutanol (δ0.88 -CH_3); as well as acrylic acid (δ5.76 -CH_2; δ6.02 -CH_2; δ6.11 -CH) and propionate (δ2.19 -CH_2; δ1.06 -CH_3). A diaper's absorbent core is predominantly derived from polymers made from very small particles of acrylic acid derivatives giving it its highly absorbent properties. Our results would suggest that either the observed contaminants are introduced during the manufacturing process or that degradation of this polymer may have occurred, which would argue for further investigation of the effect of shelf-life or storage on the chemicals that leach from diapers.

Although some of these contaminants such as polyethylene glycol in diapers[25] and propylene glycol in infant urine have been previously reported and associated with babywipes[34], to our knowledge we provide here the most comprehensive assignment of exogenous chemicals deriving from collection procedures. A systematic analysis of the effect of different diapers on synthetic urine by Goodpaster and colleagues showed that multivariate analysis of either NMR or LC-MS spectra resulted in clustering of spectra according to the brand of diaper, which is consistent with our results (Suplementary Figure 3). In Goodpaster's study, it was estimated that up to 59% of the NMR spectrum was influenced by chemical contamination from diapers. Thus, it is important that the brand of diaper is standardized across a study to prevent the introduction of extraneous variation.

Assessment of dilution factor and spectral normalization on infant urine profiles

Since urine sample dilution can vary greatly, the systematic effect of dilution on the spectral profile was assessed. The NMR spectra obtained from an individual sample, with each spectrum corresponding to a different dilution factor, is shown in Figure 3. These results indicate that after adjustment for the number of scans to account for differences in signal to noise[30], there was little visual effect of dilution on the NMR spectral signal intensities (Figure 3a). However, the dilution factor (indicated by color) was shown to affect the statistical models obtained when non-normalized spectra were used, suggesting that dilution has a physicochemical effect on the observation of certain metabolite signals, and metabolic differences between individuals were more easily

observed in the more dilute samples (indicated by symbols, Figure 3b). Once normalization was applied to the spectral data, which is typical practice for metabolic profiling studies, using either PQN normalization with reference to the median spectrum (Figure 3c) or noise (data not shown) as a standardizing factor, the effect of dilution, was reduced and clear inter-individual differences between the infants became apparent at all dilution factors, with the exception of two infants whose profiles were largely superimposed in the first two components (as indicated by the squares and stars Figure 3c). This indicates that standard normalization procedures can largely resolve the effect of dilution on the spectroscopic profiles. For one infant (represented by diamond symbols), the effect of urinary dilution was still clearly visible; this infant had the lowest creatinine to noise ratio, suggesting that the urine sample from this individual was the most dilute to begin with. A more accurate way of controlling for differential concentrations between samples would be to ensure that all samples were adjusted for osmolality prior to analysis but this would require extra sample and would introduce an extra time constraint and potential for human error. Since for three of the five infants, the two most dilute samples were situated to the left of the corresponding cluster for that individual, we assumed that an appropriate dilution factor to employ would be no more than double the original sample volume.

Age variability in infant urine profiles

Given the generally chaotic start to life, it is often difficult to standardize the exact time of biofluid sampling in infants. However this study, and many previous studies, highlights the high variability in neonatal urine composition in relation to age. Over the first few weeks of life, infant metabolism undergoes rapid change and this is reflected in the dynamic urine profiles. Here we show that urinary NMR spectral profiles of infants at the time of birth (Figure 4b) are evidently very similar to urinary samples collected from the mother at the same time (Figure 4a). Throughout pregnancy the developing fetus relies on the mother for nutrient uptake, waste removal and gas exchange; this occurs through the placenta, thus similarities in urinary global metabolic profile are observed[32,33].

The similarities between mother and newborn are particularly apparent in the aromatic region of the NMR spectra. A higher concentration of hippurate and phenylacetylglutamine (PAG) in the infant birth sample can be seen in comparison to later timepoints. We would expect to see low concentrations of phenolic and aromatic compounds in infant urine since these molecules are produced mainly by gut microbiota. We know from the microbiome literature and from studies in rodent models that colonization of the infant gut largely begins at birth and that over a period of time as the symbiotic gut bacteria evolve, there is an increase in molecules such as hippurate, PAG (or the equivalent glycine conjugate in rodents), indoles and methylamines[31].However, many papers report an absence of signals in the spectral window of infant urine where phenolic and other aromatic molecules would be expected to arise. In the urine sample obtained immediately post-partum, hippurate, 4-cresylsulfate and PAG are found in the infant urine in approximately equivalent concentrations to the maternal urine. Although there are strong similarities between mother and newborn infant urine, some differences between maternal and infant urine composition are visible at this initial sampling point, including the presence of higher concentrations of trimethylamine and formate and lower concentrations of methylhistidine.

After two weeks, the infant's urine spectrum has become more dilute compared to the mothers urine (Figure 4c). The most striking difference between these spectra is that there is much less creatinine and almost no hippurate, with a reduction in the excretion of most aromatic compounds in the week two sample from the infant compared to the birth sample. This reinforces the observation of products of mother-fetal exchange of nutrients and waste products in the infant urine samples obtained immediately post-partum. At this time, betaine is also found in higher concentrations in the neonatal urine compared to the maternal or later urine samples from the baby and is thought to reflect the as yet immature kidneys[37].

At 14 months (Figure 4d), the infant will have grown considerably, will be eating more solid foods and will have a larger community of bacteria in their gut, which contributes to metabolism and

therefore their urine metabolic profile[35]. The NMR spectra measured at 14 months showed higher concentrations of creatinine which is related to muscle mass, as well as higher levels of metabolites found in the aromatic region (e.g hippurate and PAG), which are co-metabolites produced from gut microorganisms.

Relevance of results to neonatal research

Currently most clinical tests available in neonatology rely on technologies that can only measure a single chemical in blood, urine or other biofluids; these tests are neither sensitive nor specific for any particular disease[26,27,28]. Analysis of the infant urinary metabolome offers a holistic approach to systems medicine, with the promise to enhance clinical chemistry diagnostics in several pathologic conditions related to newborns and significantly improve monitoring of physiological development and disease[29]. The goal is, for each condition or disease to derive a discrete metabolic phenotype, comprised essentially from small but significant differences in certain metabolites[10]. Thus, scientists need to be sure that these discriminant metabolites are products of a neonate's metabolism and kidney function and not just by-products of the cotton wool and diapers used to collect the biofluids.

This issue of possible contaminants has been examined and further reported using NMR-based metabonomics in three diaper brands using the 'diaper plus cotton ball' technique[30]. The authors found that the NMR spectra of the two cotton ball brands examined were identical, thus only one brand of cotton ball was used for 'diaper plus cotton ball' sample analysis. It is also indicated that although certain diaper brands are more likely to exhibit distinct contaminant profiles than others, the contamination levels of newborn urine collected using the 'diaper plus cotton ball' technique generally contain less diaper contamination than the 'diaper alone' technique. Therefore, it is advised that diapers and cotton balls are chemically characterized using metabonomic methodologies prior to analysis of study urine samples[22]. This way, contaminants arising from collection material would be manageable and unwanted variation introduced by contamination can be minimized.

Infant age is an important variable that contributes to the metabolic profile of biofluids and thus establishing the metabolic profiles linked to age has become a priority. From birth the infant urinary metabolic profile can be related to the mother and changes dramatically in the first few days after birth. The infant's body continues to develop outside of the womb, and as they start to establish a gut microbiome, the metabolic profile reiterates this through changes in their metabolome; this is particularly evident over the first few days of life. The knowledge of an age-related metabolic profile can make the interpretation of data for other effects such as disease easier, which is important in biomarker discovery for disease diagnosis or prognosis. In studies where age is not standardized, this should be appreciated and incorporated into the data modelling strategy.

Standardization of experimental design and subsequent analysis is imperative for interpretation of metabolic data generated in neonatal studies. Improved tools for metabonomic analysis incorporating a reproducible analytical pipeline for sample collection, processing and analysis, along with its integration with other 'omics' (proteomics, epigenomics etc.) analysis will make this technology valuable in the clinical setting opening new possibilities for disease diagnosis and assessment of disease risk. In particular, metabolic profiling of neonates will not only impact on diagnosis and treatment of neonatal diseases, but may deliver validated biomarkers of early life stress, which could be used for predicting individual predisposition to disease, or to the monitoring of disease progression and therapeutic management.

Recommendation Table

Factor	Recommendations
Optimize recovery of urine volumes	• Use cotton wool in diapers to collect urine • Include water and urine blank samples in analytical run, and apply database of known contaminants to identify chemical contaminants arising from diaper or cotton wool

	• Use modified centrifugation technique for extraction
Minimization of chemical contaminants	• Use same brand of collection kit (cotton wool, diapers) • Use centrifugation to minimize extraction of contaminants and maximize volume
Optimization of sampling handling	• Use one freeze-thaw cycle only • Establish clear SOP for biofluids collection and preparation
Optimization of spectral pre processing	• Normalize spectral data to the median spectrum • Remove spectral features corresponding to known chemical contaminants prior to modelling the data
Age variability	• Detailed records/metadata relating to demographic factors including age, post menstrual age and feeding • Where possible impose a narrow age range inclusion criteria.

Figure Legends

Figure 1. Schematic pipeline development for metabolic profiling of infant urine

Figure 2. a) 1D ^{1}H NMR water blank spectra from cotton wool (red spectrum) and diaper blanks (green spectrum), Key: 1. Isobutanol (δ0.88 -CH_3); 2. Propylene glycol (δ1.14 -CH_3; 3.43 -CH_2; 3.53 -CH_2; 3.87 -CH); 3. Tert-butanol (δ1.25 -CH_3); 4. Acetate (δ1.92 -CH_3); 5. N, N-dimethylacetamide (δ2.1 -CH_3; δ2.9 -CH_3; δ3.1 -CH_3); 6. Acetone (δ2.23 -CH_3); 7. Acrylic Acid (δ5.76 –CH; δ6.02 –CH; δ6.11 -CH); 8. Unknown metabolite (δ6.02); 9. Formate (δ8.46 -CH); 10. Ethylene glycol (δ3.7 –CH); 11. Propionate (δ2.19 -CH_2; δ1.06 -CH_3); b) 2D COSY ^{1}H NMR spectra, dotted line depicting Propylene glycol (δ1.14 -CH_3; 3.43 -CH_2; 3.87 -CH); c) 1D ^{1}H NMR partial urine spectra (δ1.85 - δ2.75) from cotton wool which have been subject to freeze-thaw cycles; i) urine extracted from cotton wool and then subject to one freeze thaw cycle; ii) urine extracted from cotton wool after one freeze thaw cycle; iii) urine extracted from cotton wool and acquired immediately; iv) control blank urine. Key: 1. Acetate (δ1.92 -CH_3); 2. Acetone (δ2.23 -CH_3); 3. Citrate (δ2.55 -CH_2; δ2.66 -CH_2); 4. Dimethylamine (δ2.72 -CH_3);

Figure 3. a) 1D ^{1}H NMR spectra from one infant with dilution series in effect. i) volume 1: 540µl; ii) volume 2:270 µl; iii) volume 3:135 µl; iv) volume 4: 68 µl; b) PCA of dilution series: $R^2(X)$ =0.985 (First component accounts for 84% of the total variance), $Q^2(cum)$ =0.941, c) PCA of normalized diluted infant samples: $R^2(X)$ =0.5 (First component accounts for 32% of the total variance), $Q^2(cum)$ =0.29 – colors represent dilution series; blue= volume 1: 540µl; red= volume 2:270 µl; green= volume 3:135 µl; purple= volume 4: 68 µl; shapes represent different infants

Figure 4. Mother – Baby Urine match 1D ^{1}H NMR Spectra showing the age variability from birth until 14 months postpartum. a) Mother b) Birth c) 2 weeks d) 8 weeks and e)14 months. Key: 1. Valine (δ0.99 –CH_3; δ 1.04 –CH_3); 2. Lactate (δ1.33 -CH_3; 4.11 -CH); 3. Alanine (δ1.47 -CH_3); 4. Acetate (δ 1.92 –CH_3); 5. Uromodulin (δ2.06 –CH_2) 6. Succinate (δ2.42 -CH_2); 7. Phenylacetylglutamine (PAG) (δ

2.14 $–CH_2$; δ2.46 $-CH_2$; δ3.77 $-CH_2$; δ7.35 $(–CH)_3$); 8. Citrate (δ2.55 $-CH_2$; δ2.66 –CH); 9. Creatine (δ3.02 $-CH_3$ (overlapped with creatinine); δ3.92 $-CH_2$) 10. Proline betaine (δ3.11 $-CH_3$; δ3.30 $-CH_3$); 11. Trimethylamine-N-oxide (δ3.25 $-CH_3$); 12. Glycine (δ3.55 $-CH_2$); 13. Hippurate (δ3.95 -CH); 14. N-methylnicotinamide (δ4.48 -CH); 15. Fumarate (δ6.53 $-CH_3$); 16. Urocanic acid (δ6.65); 17. Formate (δ8.45 -CH); 18. Propylene glycol (δ1.14 $-CH_3$); 19. Tert-butanol (δ1.25 $-CH_3$); 20. Acetone (δ2.24 $-CH_3$); 21. Methylguanidine* (δ2.23 $-CH_3$); 22. Trimethylamine (δ2.28 $-CH_3$); 23. Creatinine (δ3.05 $-CH_3$; δ4.06 $-CH_2$); 24. Betaine (δ3.25 $(-CH_3)_3$; δ3.89 $-CH_2$); 25. 4-cresol sulphate (δ2.22 $-CH_3$; δ7.11 –CH; δ7.18 -CH); 26. 2-oxoglutarate (δ2.42 -CH); 27. Dimethylamine (δ2.72 $-CH_3$); 28. Dimethylglycine (δ2.93 $-CH_3$); 29. 3-methyl-histidine (δ6.93 -CH; δ7.69 -CH). *tentative assignment

Supporting Information. The following files are available free of charge at ACS website http://pubs.acs.org.

Figure S1. Schematic of urine extraction

Figure S2. Urine dilution series and 1D ^{1}H NMR experimental design

Figure S3. Principal Components Analysis (PCA) scores scatter plot showing cotton ball and diaper contamination urine profiles analyzed by 1D ^{1}H NMR spectroscopy

Figure S4. Orthogonal Projections to Latent Structures Discriminant Analysis (O-PLS-DA) loadings coefficients plot showing discriminatory metabolites separating between a cotton wool balls that was squeezed vs centrifuged.

Author information

Corresponding Author

*Dr. Anisha Wijeyesekera. E-mail: anisha.wijeyesekera04@imperial.ac.uk. Tel: +44 (0)20 7594 3107.

Notes

The authors declare no competing financial interest.

Acknowledgements

The authors thank the the Imperial-National Institute for Health Research (NIHR) Clinical Phenome Centre for support. The centre is supported by the NIHR Imperial Biomedical Research Centre based at Imperial College Healthcare National Health Service (NHS) Trust and Imperial College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. F.J. acknowledges the Imperial College Stratified Medicine Graduate Training Program in Systems Medicine and Spectroscopic Profiling (STRATiGRAD) for financial support and scientific contribution. The authors would also like to thank the MRC-NIHR National Phenome Centre (NPC) which is supported by the UK Medical Research Council (MRC) [in association with National Institute for Health Research (NIHR) (England)] Grant MC_PC_12025. The financial support of Bruker Biospin, Waters Corporation, Metabometrix and Imperial College is also gratefully acknowledged by the NPC.

Abbreviations

NMR: Nuclear Magnetic Resonance, PCA: Principal Components Analysis, OPLS-DA: Orthogonal Projections to Latent Structures Discriminant Analysis, PQN: Probabilistic quotient normalization, FID: Free Induction Decay, TSP: 3-Trimethylsilyl Propionic Acid-d4 Acid Sodium Salt, PAG: Phenylacetylglutamine

References

1. Nicholson JK, Lindon JC. Systems biology: Metabonomics. *Nature*. 2008;455(7216):1054-6.
2. Nicholson JK, Connelly J, Lindon JC, Holmes E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov*. 2002;1(2):153-61.
3. Fiehn O. Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol*. 2002;48(1-2):155-71.
4. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell*. 2008;134(5):714-7.
5. Graça G, Goodfellow BJ, Barros AS, et al. UPLC-MS metabolic profiling of second trimester amniotic fluid and maternal urine and comparison with NMR spectral profiling for the identification of pregnancy disorder biomarkers. *Mol Biosyst*. **2012**;8(4):1243-54.
6. Lindon JC, Holmes E, Nicholson JK. Pattern recognition methods and applications in biomedical magnetic resonance. *Prog Nucl Magn Reson Spectrosc*. **2001**;39(1):1-40.
7. Fanos V, Antonucci R, Barberini L, Noto A, Atzori L. Clinical application of metabolomics in neonatology. *J Matern Fetal Neonatal Med*. **2012**;25 Suppl 1:104-9.
8. Fanos V, Antonucci R, Barberini L, Atzori L. Urinary metabolomics in newborns and infants. *Adv Clin Chem*. **2012**;58:193-223.
9. Fanos V, Antonucci R, Atzori L. Metabolomics in the developing infant. *Curr Opin Pediatr*. **2013**. 25(5):604-11.
10. Mussap M, Noto A, Cibecchini F, Fanos V. The importance of biomarkers in neonatology. *Semin Fetal Neonatal Med*. **2013**;18(1):56-64.
11. Fanos V, Van den Anker J, Noto A, Mussap M, Atzori L. Metabolomics in neonatology: fact or fiction? *Semin Fetal Neonatal Med*. **2013**;18(1):3-12.
12. Trump S, Laudi S, Unruh N, Goelz R, Leibfritz D. 1H-NMR metabolic profiling of human neonatal urine. *MAGMA*. **2006**;19(6):305-12.

13. Moltu SJ, Sachse D, Blakstad EW, et al. Urinary metabolite profiles in premature infants show early postnatal metabolic adaptation and maturation. *Nutrients*. **2014**;6(5):1913-30.
14. Gu H, Pan Z, Xi B, et al. 1 H NMR metabolomics study of age profiling in children. *NMR Biomed*. **2009**;22(8):826-833.
15. Foxall PJ, Bewley S, Neild GH, Rodeck CH, Nicholson JK. Analysis of fetal and neonatal urine using proton nuclear magnetic resonance spectroscopy. *Arch Dis Child - Fetal Neonatal Ed*. **1995**;73(3):F153-F157.
16. Romero R, Espinoza J, Kusanovic JP, et al. The preterm parturition syndrome. *BJOG*. **2006**;113 Suppl :17-42.
17. Dessì A, Atzori L, Noto A, et al. Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome. *J Matern Fetal Neonatal Med*. **2011**;24 Suppl 2:35-9.
18. Atzori L, Antonucci R, Barberini L, et al. 1H NMR-based metabolomic analysis of urine from preterm and term neonates. *Front Biosci (Elite Ed)*. **2011**;3:1005-12.
19. Fanos V, Caboni P, Corsello G, et al. Urinary (1)H-NMR and GC-MS metabolomics predicts early and late onset neonatal sepsis. *Early Hum Dev*. **2014**;90 Suppl 1:S78-83.
20. Fanos V, Barberini L, Antonucci R, Atzori L. Pharma-metabolomics in neonatology: is it a dream or a fact? *Curr Pharm Des*. **2012**;18(21):2996-3006.
21. Hutton NM, Schreiner RL. Urine Collection in the Neonate. *J Obstet Gynecol & Neonatal Nurs*. **1980**;9(3):165-169.
22. Farrell M, Devine K, Lancaster G, Judd B. A method comparison study to assess the reliability of urine collection pads as a means of obtaining urine specimens from non-toilet-trained children for microbiological examination. *J Adv Nurs*. **2002**;37(4):387-93.
23. Seifert JA, Ross CA, Deutsch JC, Awad JA, Norris JM. Validating methods of urine collection: a pilot study. *Urol Nurs*. **2002**;22(2):113-7.

24. Lee EJ, Arbuckle TE. Urine-sampling methods for environmental chemicals in infants and young children. *J Expo Sci Environ Epidemiol*. **2009**;19(7):625-33.

25. Goodpaster AM, Ramadas EH, Kennedy MA. Potential effect of diaper and cotton ball contamination on NMR- and LC/MS-based metabonomics studies of urine from newborn babies. *Anal Chem*. **2011**;83(3):896-902.

26. Zohouri F V, Swinbank CM, Maguire A, Moynihan PJ. Is the fluoride/creatinine ratio of a spot urine sample indicative of 24-h urinary fluoride? *Community Dent Oral Epidemiol*. **2006**;34(2):130-8.

27. Gika HG, Theodoridis GA, Wingate JE, Wilson ID. Within-day reproducibility of an HPLC-MS-based method for metabonomic analysis: application to human urine. *J Proteome Res*. **2007**;6(8):3291-303..

28. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem*. **2006**;78(13):4281-90.

29. Dona AC, Jiménez B, Schäfer H, et al. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal Chem*. **2014**;86(19):9887-94.

30. Spagou K, Wilson ID, Masson P, et al. HILIC-UPLC-MS for exploratory urinary metabolic profiling in toxicological studies. *Anal Chem*. **2011**;83(1):382-90.

31. Holmes E, Li J V, Athanasiou T, Ashrafian H, Nicholson JK. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol*. **2011**;19(7):349-59.

32. Lager S, Powell TL. Regulation of nutrient transport across the placenta. *J Pregnancy*. **2012**;2012:179827.

33. Tea I, Le Gall G, Küster A, et al. 1H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*. **2012**;7(1):e29947.

34. O'Sulivan A, He X, McNiven E et al. Metabolomic Phenotyping Validates the Infant Rhesus Monkey as a Model of Human Infant Metabolism. *JPGN.* **2013**;56: 355-363

35. Want EJ, Wilson ID, Gika H, et al. Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc*. **2010**;5(6):1005-18.

36. Palmer C, Bik EM, DiGiulio DB et al. Development of the Human Infant Intestinal Microbiota *PLoS Biol.* **2007**;5(7): e177

37. Davies SEC, Woolf DA, Chalmers RA, et al. Proton nmr studies of betaine excretion in the human neonate: consequences for choline and methyl group supply. *J. Nutr. Biochem.* **1992**; 3: 523-530

FOR TOC ONLY

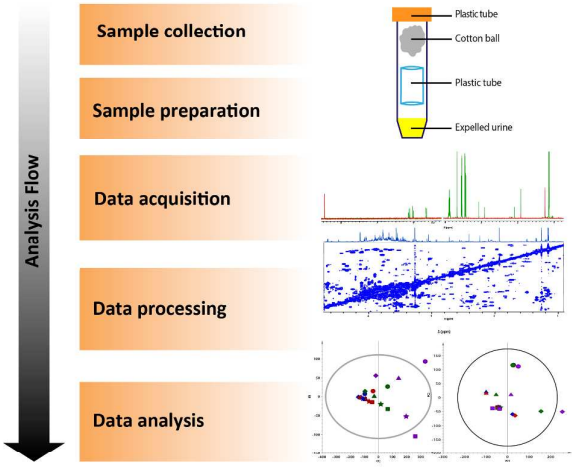
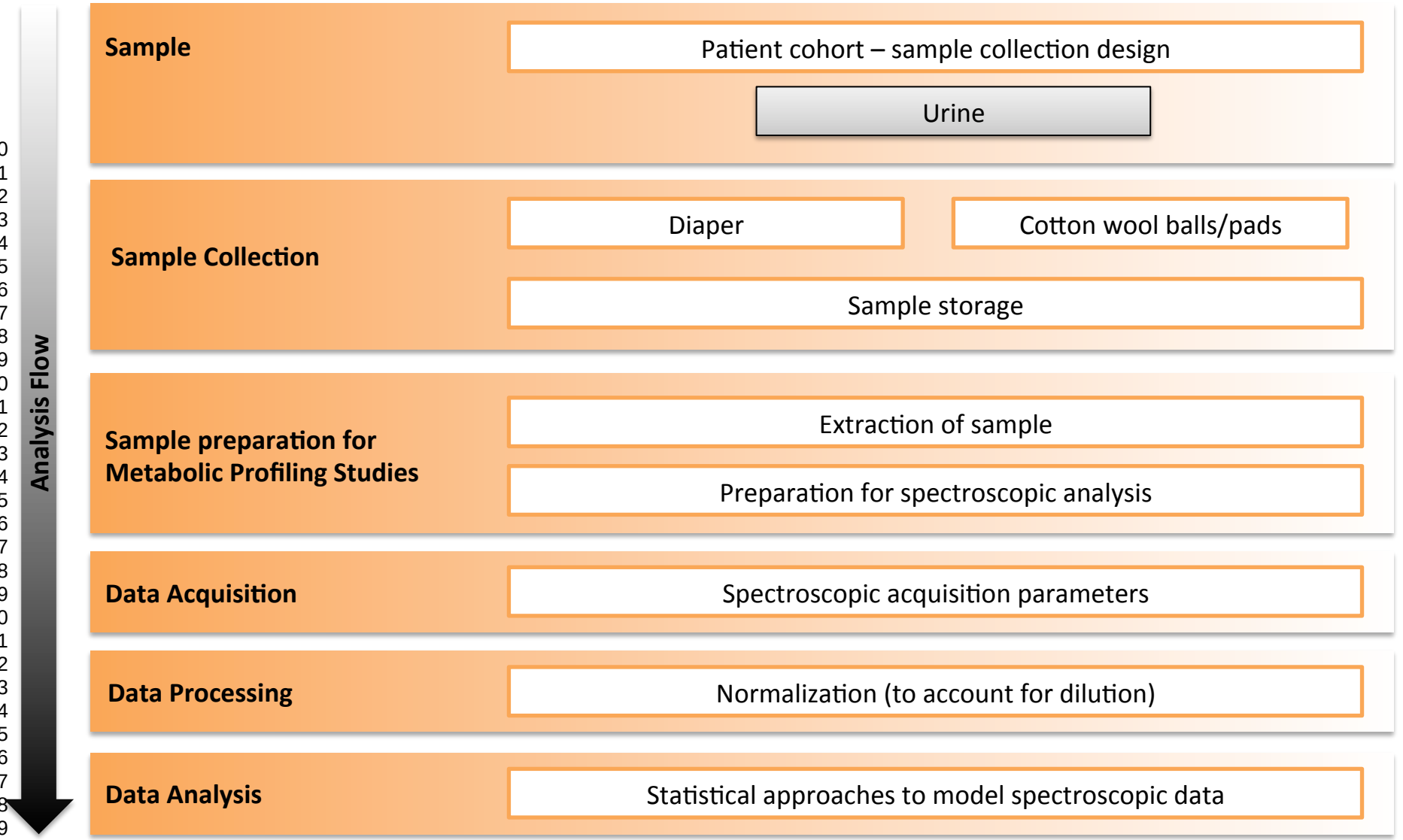

Analysis Flow
Sample collection
Sample preparation
Data acquisition
Data processing
Data analysis
Plastic tube
Cotton ball
Plastic tube
Expelled urine

Figure 1


Analysis Flow
Sample
Patient cohort – sample collection design
Urine
Sample Collection
Diaper
Cotton wool balls/pads
Sample storage
Sample preparation for Metabolic Profiling Studies
Extraction of sample
Preparation for spectroscopic analysis
Data Acquisition
Spectroscopic acquisition parameters
Data Processing
Normalization (to account for dilution)
Data Analysis
Statistical approaches to model spectroscopic data

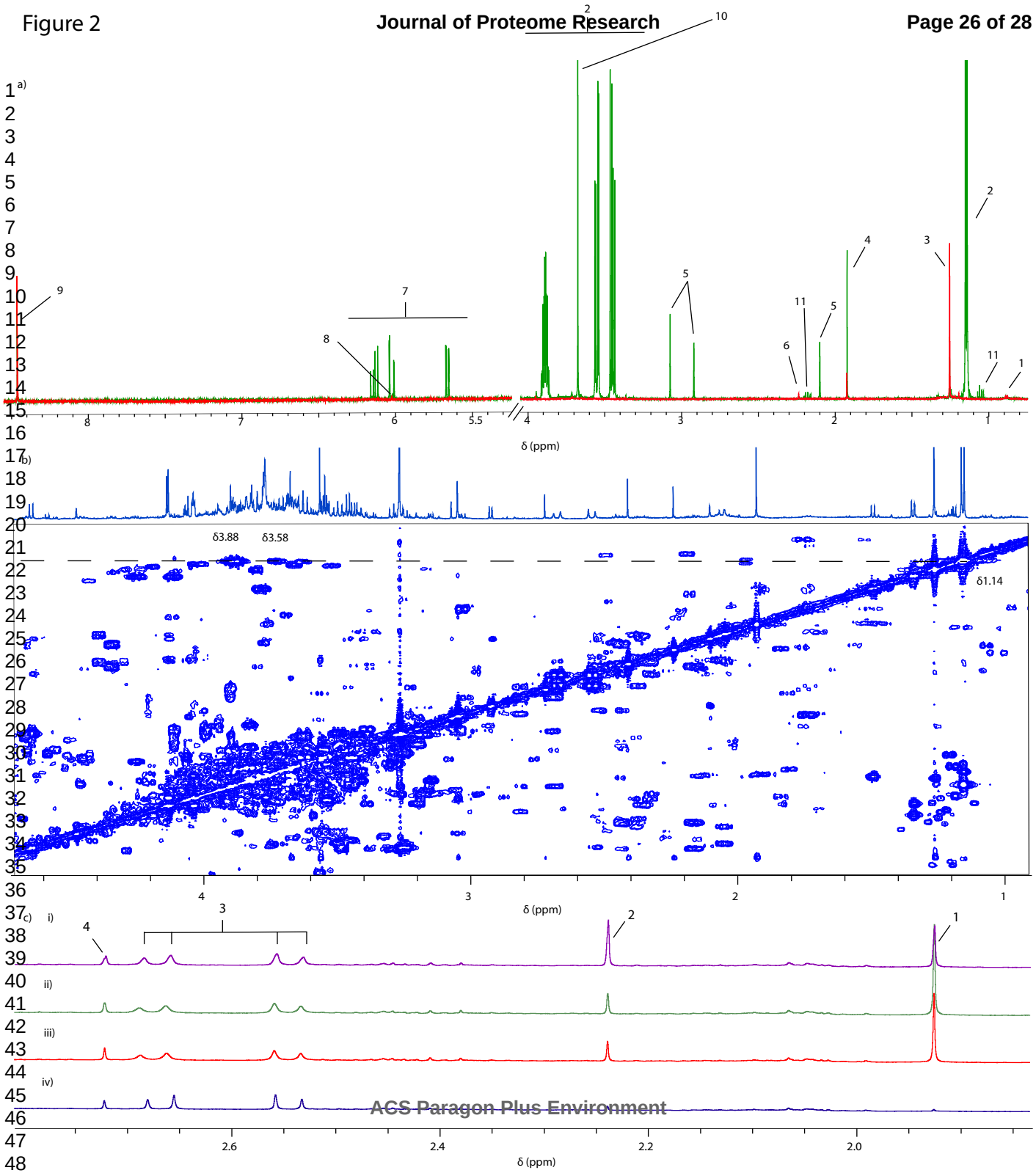

a)
b)
c)
δ (ppm)
δ3.88
δ3.58
δ1.14
i)
ii)
iii)
iv)
ACS Paragon Plus Environment

Figure 3

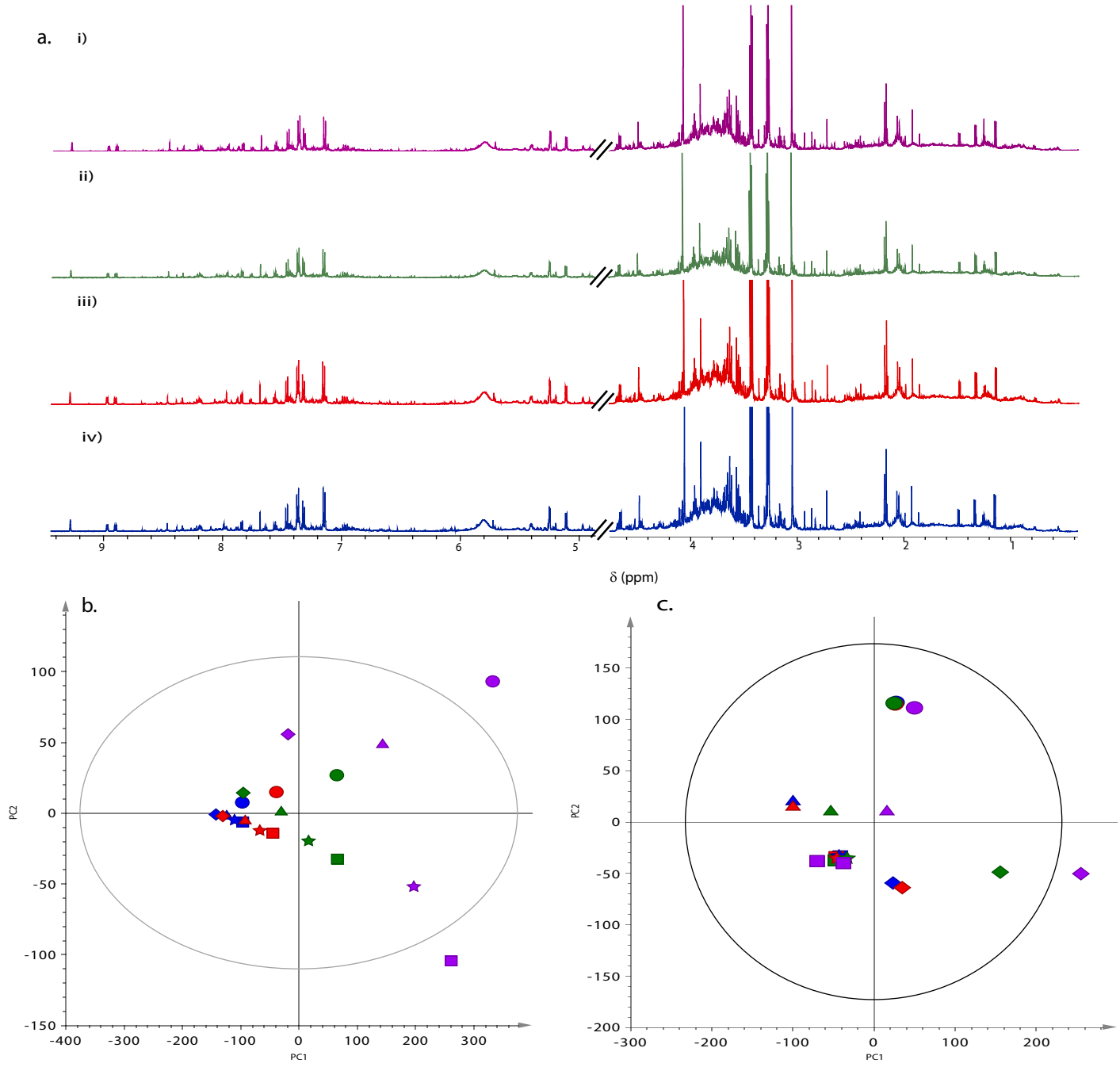
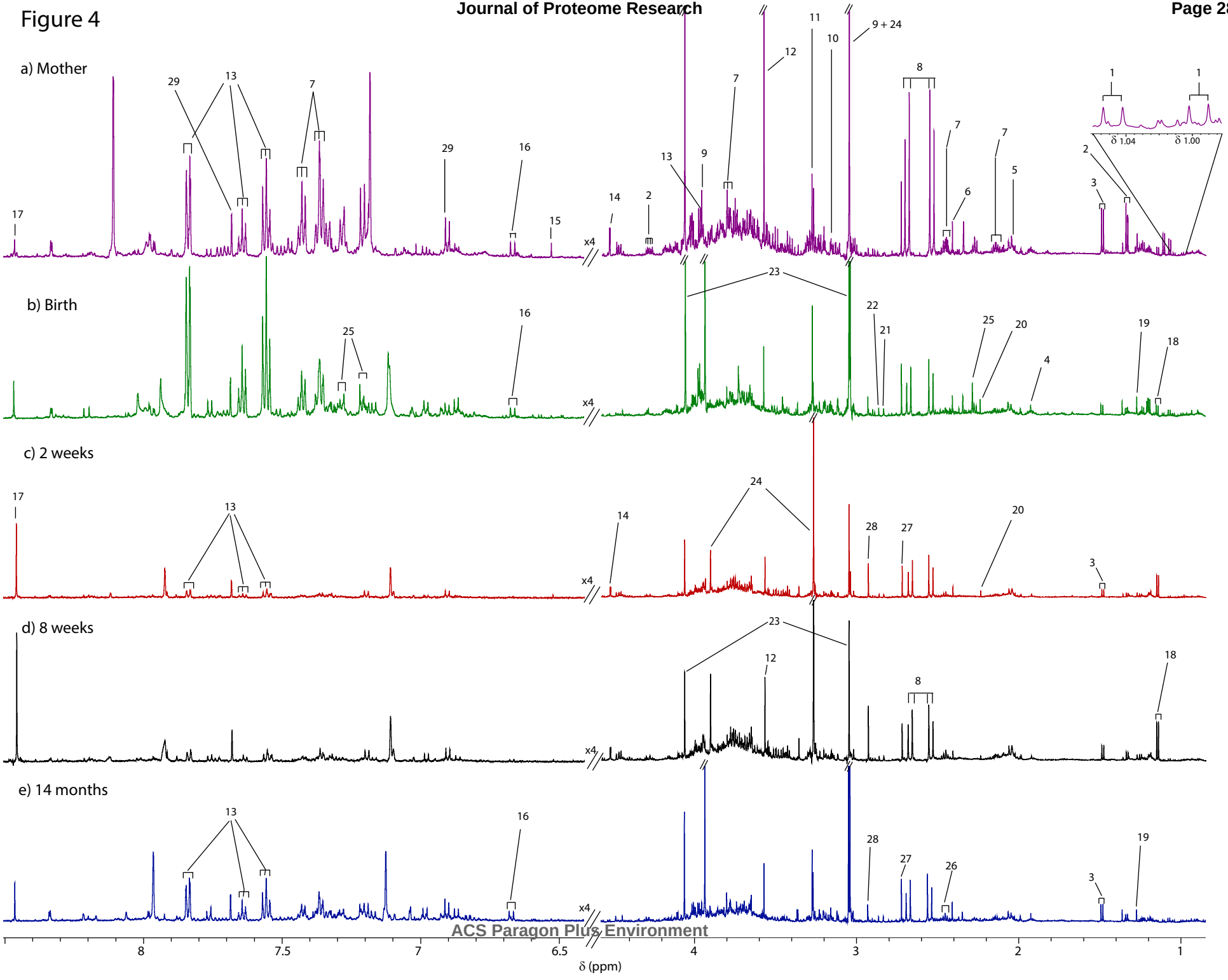


Figure 4



ACS Paragon Plus Environment