Metabolomics (2017) 13:27
DOI 10.1007/s11306-016-1153-z

ORIGINAL ARTICLE

An untargeted metabolomics method for archived newborn dried blood spots in epidemiologic studies

Lauren Petrick[1] · William Edmands[1] · Courtney Schiffman[2] · Hasmik Grigoryan[1] · Kelsi Perttula[1] · Yukiko Yano[1] · Sandrine Dudoit[2,3] · Todd Whitehead[4,5] · Catherine Metayer[4,5] · Stephen Rappaport[1,5]

Received: 13 October 2016 / Accepted: 16 December 2016 / Published online: 3 February 2017

Abstract
Introduction For pediatric diseases like childhood leukemia, a short latency period points to in-utero exposures as potentially important risk factors. Untargeted metabolomics of small molecules in archived newborn dried blood spots (DBS) offers an avenue for discovering early-life exposures that contribute to disease risks.
Objectives The purpose of this study was to develop a quantitative method for untargeted analysis of archived newborn DBS for use in an epidemiological study (California Childhood Leukemia Study, CCLS).
Methods Using experimental DBS from the blood of an adult volunteer, we optimized extraction of small molecules and integrated measurement of potassium as a proxy for blood hematocrit. We then applied this extraction method to 4.7-mm punches from 106 control DBS samples from the CCLS. Sample extracts were analyzed with liquid chromatography—high resolution mass spectrometry (LC-HRMS) and an untargeted workflow was used to screen for metabolites that discriminate population characteristics such as sex, ethnicity, and birth weight.
Results Thousands of small molecules were measured in extracts of archived DBS. Normalizing for potassium levels removed variability related to varying hematocrit across DBS punches. Of the roughly 1000 prevalent small molecules that were tested, multivariate linear regression detected significant associations with ethnicity (three metabolites) and birth weight (15 metabolites) after adjusting for multiple testing.
Conclusions This untargeted workflow can be used for analysis of small molecules in archived DBS to discover novel biomarkers, to provide insights into the initiation and progression of diseases, and to provide guidance for disease prevention.

Keywords Dried blood spots · Small molecules · LC-HRMS · Hematocrit · Metabolome

1 Introduction

Newborn dried blood spots (DBS) offer a unique resource for assessing exposures and early biological changes, years before diagnosis of clinical outcomes. DBS are collected on 'Guthrie cards' within 24–48 h of birth from >98% of the newborns in the United States to screen for inborn errors in metabolism (Gonzales 2011). Since 1982, the State of California has archived residual DBS at −20 °C for use in epidemiologic studies of pediatric outcomes (CDPH 2016). The potential utility of archived DBS for assessing in utero exposures and disease initiation is supported by targeted measurements of selected environmental contaminants and

Electronic supplementary material The online version of this article (doi:10.1007/s11306-016-1153-z) contains supplementary material, which is available to authorized users.

✉ Stephen Rappaport
srappaport@berkeley.edu

1 Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94720, USA

2 Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94720, USA

3 Department of Statistics, University of California, Berkeley, CA 94720, USA

4 Division of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720, USA

5 Center for Integrative Research on Childhood Leukemia and the Environment, University of California, Berkeley, CA 94720, USA

DNA modifications in such specimens (Funk et al. 2013; Ma et al. 2014; Wiemels et al. 1999). Furthermore, unlike most other prospective cohorts that utilize serum or plasma, DBS contain whole blood that provides access to potential biomarkers from both serum and red and white blood cells.

Untargeted metabolomics motivates a hypothesis-free approach to biomarker discovery. High-throughput analytical platforms such as liquid chromatography-high resolution mass spectrometry (LC-HRMS) can measure thousands of small molecules in a few microliters of blood. By comparing small-molecule profiles between diseased and non-diseased populations, metabolomics has been applied for discovering novel biomarkers that serve as indicators of disease processes. For example, Hazen and co-workers employed this approach to implicate joint microbial/human metabolism of the nutrient, choline, as a cause of heart disease (Koeth et al. 2013; Tang et al. 2013; Wang et al. 2011). Previous untargeted metabolomic analyses of newborn DBS have been conducted rarely, with a few examples being the studies of Denes et al. (2012), who screened for inborn errors of metabolism, and of Koulman et al. (2014) and Prentice et al. (2015) who investigated biomarkers of breastfeeding.

Methodologic validation is required before untargeted metabolomics can be applied to newborn DBS in epidemiologic studies. Measurements of small molecules in archived DBS can be affected by extraction methods, chromatography, MS ionization (Bruce et al. 2009; Contrepois et al. 2015; Michopoulos et al. 2010; Raju et al. 2016) as well as DBS aging and storage (Liu et al. 2014; Pupillo et al. 2016; Vu et al. 2011). In addition, a major concern in untargeted analysis of DBS is the effect of blood haematocrit, which can have a strong impact on quantitation when less than a whole spot is available for analysis. Hematocrit affects the spreading of a blood drop on filter paper with higher hematocrit values leading to smaller, more concentrated spots; and hematocrit can affect recoveries of small molecules and introduce matrix effects (Abu-Rabie et al. 2015; Vu et al. 2011; Youhnovski et al. 2011).

Here, we describe an untargeted metabolomics workflow for the analysis of archived DBS from the California Childhood Leukemia Study (CCLS), a population-based study containing approximately 1000 childhood leukemia cases and 1200 matched controls (Metayer et al. 2013). We measured small molecules with LC-HRMS in archived DBS from 106 control children in the CCLS. As validation of our methodology, we performed an untargeted analysis to pinpoint metabolites that were statistically associated with population characteristics. Our untargeted workflow includes: (1) data filtering and normalization for potassium levels as surrogates for hematocrit (Capiau et al. 2013; De Kesel et al. 2014), (2) multivariate statistical analysis to identify discriminating features, (3) manual examination of significant features for peak and integration quality, and (4) annotation of significant features with MSMS.

2 Materials and methods

2.1 Reagents and chemicals

Water (LC-MS Ultra Chromasolv), methanol (LC-MS Ultra Chromasolv), formic acid (eluent additive for LC-MS), acetic acid (eluent additive for LC-MS), potassium chloride (>99%), and cholic-24-^{13}C acid (>99%) were purchased from Sigma Aldrich (St. Louis, MO). Acetonitrile (Optima UHPLC-MS) and sodium chloride (>99%) were purchased from Fisher Scientific (Pittsburgh, PA). Ethanol (Koptec, 200 proof) was purchased from DLI (King of Prussia, PA). *Cis*-4,7,10,13,16,19-docosahexaenoic acid (>98%) was purchased from MP Biomedicals Inc. (Burlingame, CA).

2.2 Samples of DBS

2.2.1 Experimental DBS for method development

For method development, venous blood was collected with informed consent from an adult female volunteer in Na-heparin tubes. It was diluted or concentrated by adding or removing plasma from the same subject to produce blood with low, medium (unadjusted), or high hematocrit, as described elsewhere (Capiau et al. 2013). Experimental DBS for validation of potassium measurements were prepared by aliquoting 50 µL of whole blood on Whatman 903 Guthrie cards (Sigma Aldrich, St. Louis, MO), which were dried under vacuum for 2 weeks and stored at −20 °C prior to use (~4 months).

2.2.2 Archived newborn DBS

Archived newborn DBS for CCLS participants were obtained from the California birth registry (Sacramento, CA). We received single 4.7-mm punches (equivalent to ~8 µL whole blood) collected between 1985 and 2006 from 106 healthy control children. An additional set of 4.7-mm punches was obtained from adjacent portions of filter paper from the same Guthrie cards to use as blanks. In addition, information on child's socio-demographic characteristics was obtained from the interview with the biological parents (mainly the mother). Summary statistics can be found in the supplemental material (Table S1).

2.3 Validation of potassium measurements with an ion selective electrode

Capiau et al. (2013) extensively evaluated measurements of potassium (K^+) with a clinical chemistry analyzer as a surrogate for hematocrit in dried blood spots. In contrast to the clinical chemistry analyzer, which requires separate punches for K^+ analysis and metabolite analysis, K^+ measurements and LC-MS analysis can be performed with a single punch when a micro-ion selective electrode is used. This is particularly important for epidemiological studies, such as ours, where only a single punch is available. Thus, we used experimental DBS ("Experimental DBS for method development" section) to evaluate the reproducibility of the potassium measurements with a micro K^+ ion selective electrode (MI-442 and MI-401 1-mm tip, Microelectrodes Inc., Accumet AB250 meter, Fisher Scientific). A set of duplicate 4-mm punches was obtained from experimental DBS that had been prepared with low, medium, and high hematocrit (six punches in total). Each punch was placed in a microcentrifuge tube and extracted with 100 µL of water at room temperature for 15 min with constant agitation at 1400 rpm (ThermoMix, Eppendorf) (Capiau et al. 2013). Then, duplicate measurements of potassium were obtained for each extract with a K^+ ion selective electrode following the manufacturer's instructions (12 total K^+ measurements, considered 1 batch). The measurements were repeated an additional three times to yield four batches, with two measurements for each of the six punches within each batch (total of 48 K^+ measurements). The batch variable represents the variation in measurements associated with time. Voltage readings were converted to concentrations with a five point semi-log calibration curve for standard solutions ranging from 0.001 to 0.01 N KCl that contained 0.1 N NaCl as a potentially interfering species. For quality assurance, a single measurement of 0.05 N KCl was obtained after each batch of 12 measurements to monitor K^+ measurement drift. The measured K^+ concentrations (mean ± SD) for the three hematocrit levels were: 2.15 ± 0.20 mM (low), 3.43 ± 0.14mM (medium), and 4.95 ± 0.15 mM (high). While newborn hematocrit is typically higher than adult hematocrit, the range of K^+ measurements used here is sufficient to evaluate the reproducibility of the electrode. The following mixed-effects model was applied to evaluate the precision of K^+ measurements, after adjustment for hematocrit levels:

$$Y_{ijk} = \beta_0 + \beta_1 X_H + a_j + b_{i(j)} + e_{k(ij)}, \quad (1)$$

where: Y_{ijk} represents the potassium concentration for the k^{th} measurement ($k = 1,2$), from the i^{th} batch ($i = 1,2,3,4$), in the j^{th} punch ($j = 1,2$). In Eq. 1, the systematic variation in Y_{ijk} is defined by the sum of the intercept (β_0) and hematocrit level ($\beta_1 X_H$), and the random variation by the sum of random effects representing punch (a_j), batch ($b_{i(j)}$), and measurement ($e_{k(ij)}$). We used the marginal R^2 for linear mixed models to describe the proportion of variance explained by the fixed effects in our model 1 (Nakagawa and Schielzeth 2013). The marginal R^2 is computed as the fraction of the variance of the fixed effects to the total variance (sum of the variance of fixed effects, random effects and residuals). After fitting Eq. 1, the model had a high marginal R^2 of 0.9686, indicating that the three random effects in the model introduced negligible variance to the observed potassium measurements and that this methodology is appropriate for measuring potassium in archived DBS.

2.4 Analysis of newborn DBS

Optimal extraction of small molecules from experimental DBS was observed with 4/1 acetonitrile/water (see supplemental material, "Materials and methods" section). For analysis of newborn DBS, potassium measurements were integrated into the extraction protocol. Archived DBS punches of 4.7-mm were first extracted with 100 µL of water at room temperature for 15 min with agitation at 1400 rpm. For each batch of 24 samples, a potassium calibration curve was obtained followed by duplicate measurements of K^+ as described previously ("Validation of potassium measurements with an ion selective electrode" section). A sample calibration curve bracketing the range of observed K^+ measurements (3.33–9.94 mM) can be found in the supplemental material (Fig. S1). The robust local regression method loess (Edmands et al. 2015) was used to adjust for measurement drift within batches, most likely due to sample deposits on the electrode over repeated measurements. After measuring potassium, 400 µL of acetonitrile was added to the aqueous solution (resulting in 4/1 acetonitrile/water) and samples were agitated at 37 °C for 1 h (Incubator Genie, Scientific Industries) and stored at −20 °C for approximately 2 weeks until analysis. Immediately prior to LC-HRMS analysis, the precipitated proteins were removed by filtration (Captiva 0.2 µM, Agilent Technologies) and an internal standard was added (^{13}C-cholic acid, final concentration = 0.06 µg/mL). Extracts were analyzed with an Agilent 1290 UHPLC system connected to a 6550 QTOF HRMS (Santa Clara, US) in ESI (−) and (+) mode directly from the Captiva well.

Chromatography was carried out at 60 °C with a 0.550 mL/min flow rate on a Zorbax SB-Aq analytical column (1.8 µM, 2.1 × 50 mM) with a Zorbaq-SB-C8 guard column (3.5 µM, 2.1 × 30 mM) with the following solvents: Buffer A: 0.2% acetic acid in water, and Buffer B: 0.2% acetic acid in methanol. A 19-min gradient was used (2–98% B in 13 min, hold at 98% B for 6 min), followed by a 3-min column re-equilibration phase. Samples were kept at 4 °C

in the autosampler during analysis. The total sample injection volume was 10 µL consisting of alternating volumes of water and sample for a total volume of 20 µL injected onto the column. Full mass spectra were acquired at 1.67 spectra/s in the range 50–1000 *m/z* (Supplemental Material, "Results and discussion" section). To monitor system stability, a pooled QC sample—prepared by combining aliquots from all DBS extracts—was injected after each group of 10 samples. Due to a LC sampler thermostat malfunction during the ESI (+) mode data collection, only the ESI (−) data was used for further analysis.

2.5 Data processing

Raw data were converted to mzXML format for peak picking using proteoWizard software (Spielberg Family Center for Applied Proteomics, Los Angeles, CA). Peak detection and retention time alignment were performed with the XCMS package (Patti et al. 2012; Smith et al. 2006) within the R statistical programming environment (version 3.2.2, R Core Team 2015). Parameters for peak picking included centwave feature detection, orbiwarp retention time correction, minimum fraction of samples in one group to be a valid group = 0.25, isotopic ppm error = 10, width of overlapping *m/z* slices (mzwid) = 0.015, retention time window (bw) = 2 s, minimum peak width = 2 s, and maximum peak width = 20 s. Peaks were grouped and filled ('group' and 'fill' function in XCMS), and the resulting peak tables of retention times, *m/z* values, and peak areas were exported for further processing. Subsequent analyses were also performed with the R platform.

2.6 Statistical analyses

Exported features from LC-HRMS analysis of newborn DBS were log-transformed and normalized prior to statistical analysis. First, the following multivariate linear regression models (Eq. 2) were used to adjust the intensity of each tested feature for the order of analysis, which includes autosampler plate number and run order, and for the potassium concentration as a surrogate for hematocrit (note that the continuous run order variable is nested within the dichotomous plate variables):

$$Y_i = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_3 (P_1 \times R) + \beta_4 (P_2 \times R) + \beta_5 K + \varepsilon_i, \quad (2)$$

where Y_i is a vector of logged feature intensities for the i^{th} feature (length = 106), P_1 and P_2 are binary variables indicating plate number [$P_j = 1$ if sample is in plate j and 0 otherwise ($j = 1,2$)], R is a numeric vector of run order, and K is a numeric vector of K^+ concentrations. Then, residuals from Eq. 2 were full-quantile normalized, as implemented in the limma R package, to make the distributions of the features comparable across all subjects (Bolstad et al. 2003; Ritchie et al. 2015).

The following multivariate linear regression model was fitted to the normalized residuals to find associations between each feature's intensity and the subjects' birth weight, sex, ethnicity, and DBS-age in years:

$$Y_i = \beta_0 + \beta_1 X_{sex} + \beta_2 X_{birth\ \ weigh_t} + \beta_3 X_{ethnicity} + \beta_4 X_{DBS-age} + \varepsilon_i, \quad (3)$$

where Y_i is a vector of logged intensities for the i^{th} feature, X_{sex} (0 = male, 1 = female) and $X_{ethnicity}$ (0 = Hispanic, 1 = non-Hispanic) are binary vectors, and $X_{birth\ weight}$ and $X_{DBS\text{-}age}$ (2015-child birth year) are numeric vectors.

Three subjects were removed from the statistical analysis due to missing covariate data, resulting in a total of 103 subjects. Since ethnicity and sex are dichotomous variables, birth weight and DBS-age variables were standardized by subtracting the mean and dividing by the standard deviation. The DBS-age covariate was included to adjust for any changes in feature intensity due to storage conditions (Koulman et al. 2014; Pupillo et al. 2016). Estimated p values obtained from the regression model were used to evaluate whether each coefficient was significantly associated with the feature abundance. These p values should be interpreted with care since they depend on several parametric assumptions, which were not evaluated. Significance levels were adjusted for multiple testing using the Benjamini-Hochberg (BH) procedure to control the false discovery rate (FDR) at $\alpha = 0.05$ (Benjamini and Hochberg 1995).

Permutation tests were then used as further, non-parametric tests of significance for the same associations (Anderson and Braak 2003). For each of the features, the same multivariate regression as in Eq. 3 was used to obtain coefficients of the covariates for the original data set. Then, for each feature, values for a given covariate of interest (e.g. ethnicity or birth weight) were permuted among the 103 subjects keeping all else constant, and fit to the model (Eq. 3). The coefficient estimate for the tested covariate was recorded for 5000 permutation iterations. The permutation p value for the covariate's coefficient was calculated as the proportion of the 5000 coefficients that were greater than the original coefficient in absolute value. This permutation test was performed for both ethnicity and birth weight, and across all features. Significance was determined after correction for FDR using the BH method.

3 Results and discussion

3.1 Detection of metabolites

LC-HRMS analysis of the extracts from newborn DBS detected 66,096 features in ESI (−) mode. Filtering for

features that were present in at least 75% of all DBS, with a mean fold change >3 compared to filter-paper extracts, and with a CV <25% in the pooled-QC injections, reduced the number of features to 3157. Additional pre-processing was performed with MetMSLine software (Edmands et al. 2015) to impute any remaining zero values with half the minimum non-zero feature abundance, to remove outliers by PCA, based on a tolerance of 95% beyond the Hotellings T2 ellipse, and to eliminate redundant signals due to in-source fragmentation and product formation (Broeckling et al. 2014). This reduced the total number of testable features to 1107.

3.2 Normalization for run order and potassium

Metabolomic data require normalization to remove unwanted systematic biases so that only biologically relevant differences are present. It is expected that higher hematocrit levels would lead to a positive bias in archived DBS measurements. Thus, there should be a positive relationship between total usable signal (TUS), i.e. the sum of all feature peak integrals per subject, and hematocrit. Here, we measured K^+ concentrations of DBS extracts as proxies for hematocrit (Capiau et al. 2013; De Kesel et al. 2014). As can be seen in Fig. 1a, there is positive correlation (Pearson correlation coefficient 0.45) between the potassium level in a DBS extract, represented by the K^+ concentration, and TUS. Adjusting each feature by only the analysis order had a modest effect on reducing this correlation (Fig. 1b, Pearson correlation coefficient 0.38). However, adjusting for both analytical order and K^+ concentration (Eq. 2) eliminated the correlation between TUS and K^+ (Fig. 1c, Pearson correlation coefficient 0.02).

As TUS was moderately correlated with K^+ and requires no additional measurements, it is an attractive alternative to using potassium. However, TUS normalization is less precise than K^+ normalization and sensitive to outliers. Furthermore, when K^+ is replaced by logTUS (Eq. 2), not all unwanted K^+ variation is removed (Fig. S2, Pearson's correlation coefficient 0.112). Therefore, we caution its use without a full evaluation of its applicability.

Potassium measurements from stored DBS and in stored extracts were found to remain stable for several months even at room temperature (Capiau et al. 2013; den Burger et al. 2015). In addition, no correlation was observed between K^+ measurements and 'age of spot' (Pearson correlation coefficient −0.085), indicating that potassium measurements were not associated with storage conditions.

3.3 Fetal metabolome coverage

We first performed qualitative analyses to characterize the fetal blood metabolome. Small molecules were annotated

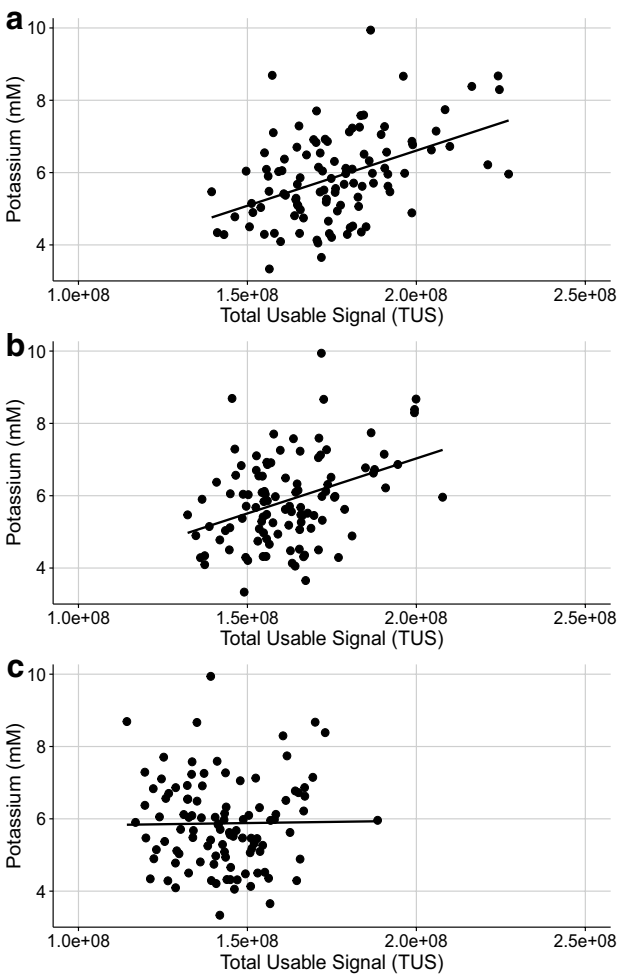

Fig. 1 Plots of potassium concentration versus total usable signal for each subject. **a** no adjustment; **b** adjustment for analytical run order; and **c** adjustment for run order and potassium concentration

using the compMS2Miner package (Edmands 2016) by comparing accurate mass and MSMS fragmentation patterns with the Human Metabolome Database (HMDB) (Wishart et al. 2013) and METLIN (Smith et al. 2005). The correlation network in Fig. 2a summarizes the annotations, where nodes are the metabolites and edges are the correlations between their peak integrals (Shannon et al. 2003). Metabolites that cluster together indicate possible shared metabolic pathways. Using a Pearson correlation cut-off of >0.7, 3960 edges were drawn. Of the 1107 metabolites, 65% of the features with MSMS spectral data ($n = 176$) were annotated by at least compound class, which included phospholipids, hormones and steroids, and saturated and unsaturated fatty acids. Steroid concentrations in newborns can be as low as 10 ng/mL (Kim et al. 2015), while polyunsaturated fatty acid concentrations (PUFA) in cord blood can be as high as 1 mg/mL (Niinivirta et al. 2011), highlighting the dynamic range of the method.

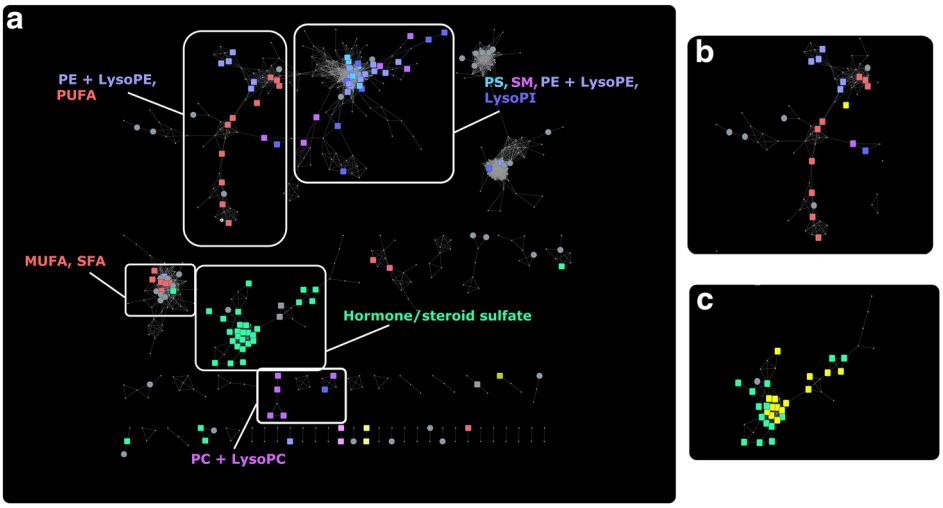

Fig. 2 **a** Correlation network of 1107 fetal blood metabolites made in Cytoscape with uncorrelated metabolites removed for clarity. *Legend* phosphatidylserine (PS), phosphatidylcholine (PC), sphingomyelin (SM), phosphatidylethanolamine (PE), phosphatidylinositol (PI), mono-unsaturated fatty acid (MUFA), poly-unsaturated fatty acid (PUFA), saturated fatty acid (SFA). **b** Expanded view of PUFA cluster. Metabolite #13338 clustered with PUFAs (#13338 highlighted in *yellow*). **c** Expanded view of hormone cluster. Features significantly associated with birth weight clustered with hormone/steroid sulfates (significant features highlighted in *yellow*)

3.4 Associations with ethnicity and birth weight

To assess our ability to discriminate biologically relevant small molecules, we screened for differences associated with the population characteristics: sex, ethnicity, and birth weight using multivariate regression models (Eq. 3) for each tested metabolite ($n = 1107$).

Initially, 19 discriminating small molecules for ethnicity and birth weight were determined by their p-values after BH correction. Peak morphology and integration quality were assessed manually for each discriminating metabolite, and resulted in exclusion of 1 feature. The curated list of 18 discriminating metabolites is summarized in Table 1. All small molecules with statistically significant birth weight and ethnicity coefficients from the multivariate regression models also had significant coefficients for these variables under the permutation tests, providing further evidence that there is a significant association between birth weight or ethnicity and feature intensity for these features.

The "DBS age" covariate was not significantly associated with any of the discriminating molecules (data not shown), suggesting that these metabolites were not affected by the duration of DBS storage at −20 °C. This is particularly important because there was an overall affect of spot age on DBS extraction efficiency (e.g. older spots extracted less efficiently). Prior to data normalization there was a weakly negative correlation between TUS and DBS age (Pearson's correlation coefficient −0.17), which was enhanced after adjusting for run order and potassium (Pearson's correlation coefficient = −0.28). This indicates that the age of spot affected extraction efficiency, and further validates the use of DBS age as a covariate in the statistical model (Eq. 3).

Table 1 Selected coefficients from linear model (Eq. 3) for discriminating features using logged levels of the metabolites

Metabolite no.	Coefficient	Adj. p value	Permutation adj. p value
Ethnicity			
13,338	0.305	0.0277	<0.0001
38,341	−0.339	0.0432	<0.0001
14,374	0.235	0.0475	<0.0001
Birth weight			
22,415	−0.176	0.0383	0.0260
18,138	−0.159	0.0337	0.0158
32,014	−0.205	0.0116	<0.0001
19,412	−0.226	0.0116	<0.0001
19,411	−0.230	0.0178	<0.0001
22,424	−0.156	0.0383	0.0158
17,929	−0.182	0.0178	0.0158
16,538	−0.214	0.0116	<0.0001
27,970	−0.130	0.0425	0.0260
17,919	−0.317	0.0132	<0.0001
19,409	−0.280	0.0116	<0.0001
29,345	−0.171	0.0337	0.0158
17,918	−0.322	0.0178	<0.0001
19,413	−0.255	0.0116	<0.0001
16,685	−0.226	0.0116	<0.0001

Table 2 Summary of acquired MS experimental data and annotations of discriminating metabolite features

Metabolite no.	*m/z*	Retention time (sec)	Prominent MS/MS fragments	Fragment ID	Putative ID	Species(-)	MSI confidence level	Molecular formula	Delta ppm
17,918	383.1165	235	59.0158, 164.9403, 303.1594	Sulfate neutral loss	Unknown hormone sulfate		3		
19,409	399.1481	244	96.9606	Sulfate	Unknown hormone sulfate		3		
19,413	399.1479	276	96.9606	Sulfate	Unknown hormone sulfate		3		
17,919	383.1164	285	59.0157, 164.9397, 303.1599	Sulfate neutral loss	Unknown hormone sulfate		3		
19,411	399.1480	289	96.9606	Sulfate	Unknown hormone sulfate		3		
19,412	399.1479	318	96.9606	Sulfate	Unknown hormone sulfate		3		
22,415	429.1949	355	96.9606	Sulfate	Conjugated androgen	M + Hac-H	3	$C_{19}H_{30}O_5S$	−0.806
18,138	385.1688	368	96.9606	Sulfate	Conjugated androgen	M-H	3	$C_{19}H_{30}O_6S$	−0.613
17,929	383.1535	401	96.9606	Sulfate	Conjugated androgen	M-H	3	$C_{19}H_{28}O_6S$	0.305
22,424	429.1948	404	59.0156, 96.9606	Sulfate	Conjugated androgen	M + Hac-H	3	$C_{19}H_{30}O_5S$	−1.039
16,538	367.1217	409	59.0157,96.9604	Sulfate	Unknown hormone sulfate		3		
16,685	369.1370	416	59.0159, 96.9604	Sulfate	Unknown hormone sulfate		3		
27,970	483.0777	453	none				4		
32,014	525.2522	609	59.0158, 96.9605, 189.9560	Sulfate	Unknown hormone sulfate		3		
14,374	341.2108	671	none				4		
29,345	497.2938	675	59.0157,96.9608,158.9386, 303.2320	Sulfate	Unknown hormone sulfate		3		
38,341	595.2878	712	78.9592,152.9957, 241.0113,279.2322	Inositol headgroup fragments	Unknown PI		3		
13,338	327.2329	712	59.0156,283.2420, 229.1958	Alkyl chain	Docosahexaenoic acid	M-H	1	$C_{22}H_{32}O_2$	−0.138

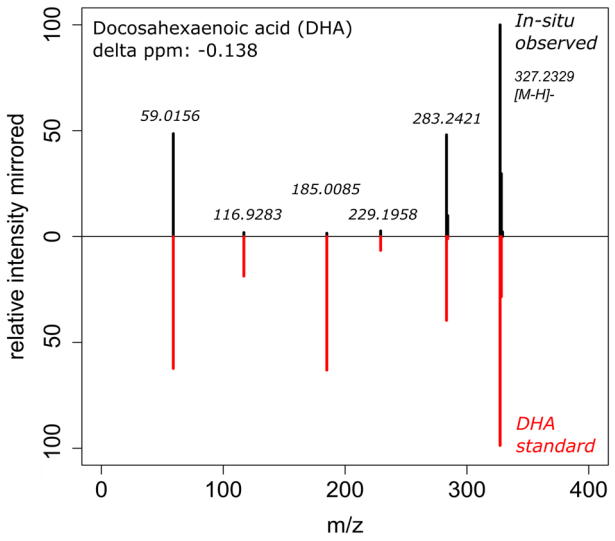

Fig. 3 Mirror plot of observed MSMS for feature #13338 (*top*) and MSMS of a commercially purchased DHA standard analyzed under the same experimental parameters (*bottom*)

3.5 Annotation of discriminating features

For further validation of the methodology, several of the significant features from "Associations with ethnicity and birth weight" section were putatively annotated to ensure biological plausibility. As shown in Table 2, the CompMS2Miner package facilitated the annotation of 16 discriminating features based on accurate mass and tandem mass spectra. For example, metabolite #13338 clustered with PUFA in the correlation network (Fig. 2b) and had an accurate mass of *m/z* 327.2329 within −0.138 ppm of docosahexaenoic acid (DHA, $[M\text{-}H]^-$). When compared to a commercial DHA standard, the observed peak eluted within 2 s of the standard peak and showed excellent MSMS spectral match between major fragments of *m/z* 283.2421, 229.1958, 185.0054 and 59.0156 (Fig. 3). As major nutritional sources of DHA are fish and fish oil supplements, it is possible that the higher levels of DHA observed in the non-Hispanic population is due to dietary differences associated with income status. In fact, 62% of

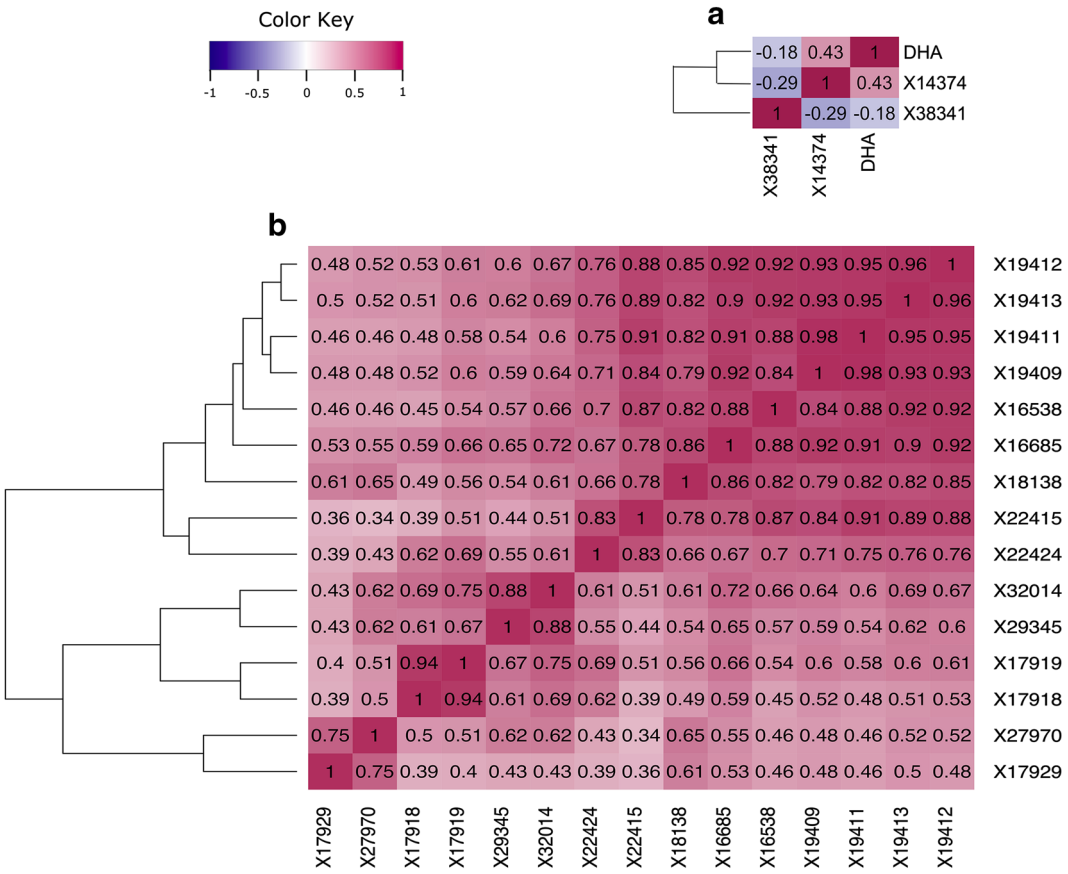

Fig. 4 Correlation matrices and dendrogram of discriminating features for ethnicity (**a**) and birth weight (**b**) based on Pearson correlation and ordered using agglomerative hierarchical clustering calculated from complete linkage using Euclidean distance ('hclust' function in R)

the Hispanic subjects were lower income (defined as annual household income <$45,000/yr), as opposed to only 10% of the non-Hispanic group. DHA consumption has been found to be lower in low-income women during pregnancy and lactation (Nochera et al. 2011).

Although the identity of metabolite #14374 is unknown, it was moderately correlated with DHA ($r=0.43$, Fig. 4), eluted at a similar retention time (within 45 s), and was also present at higher abundance in the non-Hispanic population. Feature #38341 was not correlated with the other distinguishing metabolites for ethnicity, was higher in the Hispanic population, and had multiple fragments characteristic of the inositol head group of a phosphatidylinositol at *m/z* 152.9957 and 241.0113 (Hsu and Turk 2000).

Most of the features that discriminated for birth weight contained an MSMS base fragment ion at *m/z* 96.9606, characteristic of a sulfate group. Even in the absence of distinguishable MSMS fragments to allow an exact annotation, these features had correlation coefficients ranging from 0.32 to 0.98, and clustered in the correlation network with endogenous steroids and hormones (Figs. 2c, 4b). Two sets of isomeric metabolites were observed with accurate masses of *m/z* 399.148 (#19409, #19411, #19412, #19413) and 429.194 (#22415, #22424), which were all highly correlated with each other and with metabolites #16685, #16535, and #18138 (correlation coefficient ranging from 0.70 to 0.98). Based on comparison with exact mass, metabolites #18138, #22424, and #22415 were putatively identified as conjugated androgen steroids. A third set of isomeric features observed at *m/z* 383.116 (#17918, #17919) were highly correlated with each other ($r=0.94$), but only moderately correlated with the other distinguishing features for birth weight. Metabolite #17929 was putatively identified as a conjugated androgen.

Several of the discriminating metabolites had accurate masses similar to conjugated forms of steroids previously measured in human blood or amniotic fluid. For example, 16a-hydroxyDHEAS ($C_{19}H_{28}O_6S$, #17929) has been measured in arterial cord plasma (Mitchell and Shackleton 1969), 5-androstrenetriol (sulfated form, $C_{19}H_{30}O_6S$, #18138) has been measured in amniotic fluid immediately prior to delivery (Schindler and Siiteri 1968), and androsterone sulfate ($C_{19}H_{30}O_5S$, #22415, #22424) and 16a-hydroxyDHEAS ($C_{19}H_{28}O_6S$, #17929) have been measured in infant plasma (Sánchez-Guijo et al. 2015).

All of the discriminating features were negatively associated with birth weight. Although several steroids including progesterone, 17-hydroxyprogesterone (17-OHP), cortisol, and growth hormones have been found to be inversely associated with infant birth weight, newborn hormone levels are influenced by many factors such as delivery type, gestational age, maternal height, maternal hormone levels, and the day of sample collection (Carlsen et al. 2006; Lagiou et al. 2014; Rajesh et al. 2000; Schwarz et al. 2009).

4 Conclusions

We developed an LC-HRMS method for interrogating the fetal metabolome from archived newborn DBS (4.7-mm punches). Control punches of DBS from the CCLS, equivalent to ~8 µL of whole blood, were used for validation. Over 1000 prevalent and robust features were measured in the blood extracts, representing all of the major subclasses of metabolites. By measuring potassium in each punch as a proxy for hematocrit, we were able to adjust for nuisance technical variation. Even with a small sample size of 103 newborn DBS, we were able to identify 18 biologically plausible features that could discriminate for newborn birth weight and ethnicity. We are currently applying this methodology to newborn DBS from the CCLS, to seek features that discriminate between childhood leukemia cases and controls.

Acknowledgements We gratefully acknowledge the assistance of Agilent Technologies (Santa Clara, CA, USA) for the loans of the liquid-chromatography mass-spectrometry instruments used in these analyses. We also thank the families for their participation in the CCLS.

Funding This work was supported by the National Institute for Environmental Health Sciences of the U.S. National Institutes of Health (NIEHS) and the U.S. Environmental Protection Agency through grants to the Center for Integrative Research on Childhood Leukemia and the Environment (NIEHS grants P01 ES018172 and P50ES018172 and USEPA grants RD83451101 and RD83615901), by the California Childhood Leukemia Study (NIEHS grants R01ES009137 and P42ES004705), by NIEHS grant P42ES0470518, and by a post-doctoral fellowship from the Environment and Health Fund, Jerusalem, Israel.

Compliance with ethical standards

Conflict of interest L. Petrick, W. Edmands, C. Schiffman, H. Grigoryan, K. Perttula, Y. Yano, S. Dudoit, T. Whitehead, C. Metayer and S. Rappaport have no conflict of interest to declare.

Disclaimer The ideas and opinions expressed herein are those of the authors and do not necessarily represent the official views of EPA or NIEHS. Endorsement of any product or service mentioned is not intended nor should it be inferred.

Ethics approval The study was approved by the University of California Committee for the Protection of Human Subjects, the California Health and Human Services Agency Committee for the Protection of Human Subjects, and the institutional review boards of all participating hospitals.

Informed consent Written informed consent was obtained from the adult volunteer subject and parents of all participating subjects.

References

Abu-Rabie, P., Denniff, P., Spooner, N., Chowdhry, B. Z., & Pullen, F. S. (2015). Investigation of different approaches to incorporating internal standard in DBS quantitative bioanalytical workflows and their effect on nullifying hematocrit-based assay bias. *Analytical Chemistry, 87*(9), 4996–5003. doi:10.1021/acs.analchem.5b00908.

Anderson, M., & Braak, C. Ter (2003). Permutation tests for multifactorial analysis of variance. *Journal of Statistical Computation and Simulation, 73*(2), 85–113. doi:10.1080/00949650215733.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, 57*(1), 289–300. doi:10.2307/2346101.

Bolstad, B. M., Irizarry, R., Astrand, M., & Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics (Oxford, England), 19*(2), 185–193. doi:10.1093/bioinformatics/19.2.185.

Broeckling, C. D., Afsar, F. A., Neumann, S., Ben-Hur, A., & Prenni, J. E. (2014). RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Analytical Chemistry, 86*(14), 6812–6817. doi:10.1021/ac501530d.

Bruce, S. J., Tavazzi, I., Parisod, V., Rezzi, S., Kochhar, S., & Guy, P. a (2009). Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Analytical Chemistry, 81*(9), 3285–3296. doi:10.1021/ac8024569.

California Department of Public Health (CDPH). (2016). Background and History of the California Biobank Program (CBP). https://www.cdph.ca.gov/programs/GDSP/Pages/MoreAboutTheCBP.aspx. Accessed 8 Aug 2016.

Capiau, S., Stove, V. V., Lambert, W. E., & Stove, C. P. (2013). Prediction of the hematocrit of dried blood spots via potassium measurement on a routine clinical chemistry analyzer. *Analytical Chemistry, 85*(1), 404–410. doi:10.1021/ac303014b.

Carlsen, S. M., Jacobsen, G., & Romundstad, P. (2006). Maternal testosterone levels during pregnancy are associated with offspring size at birth. *European. Journal of Endocrinology: European Federation of Endocrine Societies*, 155(2), 365–370. doi:10.1530/eje.1.02200.

Contrepois, K., Jiang, L., & Snyder, M. (2015). Optimized analytical procedures for the untargeted metabolomic profiling of human urine and plasma by combining hydrophilic interaction (HILIC) and reverse-phase liquid chromatography (RPLC)–mass spectrometry. *Molecular & Cellular Proteomics, 14*(6), 1684–1695. doi:10.1074/mcp.M114.046508.

De Kesel, P. M. M., Capiau, S., Stove, V. V., Lambert, W. E., & Stove, C. P. (2014). Potassium-based algorithm allows correction for the hematocrit bias in quantitative analysis of caffeine and its major metabolite in dried blood spots. *Analytical and Bioanalytical Chemistry, 406*(26), 6749–6755. doi:10.1007/s00216-014-8114-z.

den Burger, J.C.G., Wilhelm, A. J., Chahbouni, A. C., Vos, R. M., Sinjewel, A., & Swart, E. L. (2015). Haematocrit corrected analysis of creatinine in dried blood spots through potassium measurement. *Analytical and Bioanalytical Chemistry, 407*(2), 621–627.

Dénes, J., Szabó, E., Robinette, S. L., Szatmári, I., Szőnyi, L., Kreuder, J. G., et al. (2012). Metabonomics of newborn screening dried blood spot samples: A novel approach in the screening and diagnostics of inborn errors of metabolism. *Analytical Chemistry, 84*(22), 10113–10120. doi:10.1021/ac302527m.

Edmands, W. M. (2016). CompMS2miner: a metabolite identification R package. https://github.com/WMBEdmands/CompMS2miner. doi:10.5281/zenodo.56582. Accessed 29 September 2016.

Edmands, W. M., Barupal, D. K., & Scalbert, A. (2015). MetMSLine: An automated and fully integrated pipeline for rapid processing of high-resolution LC-MS metabolomic datasets. *Bioinformatics (Oxford, England), 31*(5), 788–790. doi:10.1093/bioinformatics/btu705.

Funk, W. E., McGee, J. K., Olshan, A. F., & Ghio, A. J. (2013). Quantification of arsenic, lead, mercury and cadmium in newborn dried blood spots. *Biomarkers: Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals, 18*(2), 174–177. doi:10.3109/1354750X.2012.750379.

Gonzales, J. L. (2011). Ethics for the pediatrician: Genetic testing and newborn screening. *Pediatrics in Review, 32*(11), 490–493. doi:10.1542/pir.32-11-490.

Hsu, F. F., & Turk, J. (2000). Characterization of phosphatidylinositol, phosphatidylinositol-4-phosphate, and phosphatidylinositol-4,5-bisphosphate by electrospray ionization tandem mass spectrometry: A mechanistic study. *Journal of the American Society for Mass Spectrometry, 11*(11), 986–999. doi:10.1016/S1044-0305(00)00172-0.

Kim, B., Lee, M. N., Park, H. D., Kim, J. W., Chang, Y. S., Park, W. S., & Lee, S. Y. (2015). Dried blood spot testing for seven steroids using liquid chromatography-tandem mass spectrometry with reference interval determination in the Korean population. *Annals of Laboratory Medicine, 35*(6), 578–585. doi:10.3343/alm.2015.35.6.578.

Koeth, R. A., Wang, Z., Levison, B. S., Buffa, J. A., Org, E., Sheehy, B. T., et al. (2013). Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature Medicine, 19*, 576–585. doi:10.1038/nm.3145.

Koulman, A., Prentice, P., Wong, M. C. Y., Matthews, L., Bond, N. J., Eiden, M., et al. (2014). The development and validation of a fast and robust dried blood spot based lipid profiling method to study infant metabolism. Metabolomics, 1–8. doi:10.1007/s11306-014-0628-z.

Lagiou, P., Samoli, E., Hsieh, C. C., Lagiou, A., Xu, B., Yu, G. P., et al. (2014). Maternal and cord blood hormones in relation to birth size. *European Journal of Epidemiology, 29*(5), 343–351. doi:10.1007/s10654-014-9914-3.

Liu, G., Mühlhäusler, B. S., & Gibson, R. A. (2014). A method for long term stabilisation of long chain polyunsaturated fatty acids in dried blood spots and its clinical application. *Prostaglandins Leukotrienes and Essential Fatty Acids, 91*(6), 251–260. doi:10.1016/j.plefa.2014.09.009.

Ma, W.-L., Gao, C., Bell, E. M., Druschel, C. M., Caggana, M., Aldous, K. M., et al. (2014). Analysis of polychlorinated biphenyls and organochlorine pesticides in archived dried blood spots and its application to track temporal trends of environmental chemicals in newborns. *Environmental Research, 133*, 204–210. doi:10.1016/j.envres.2014.05.029.

Metayer, C., Zhang, L., Wiemels, J. L., Bartley, K., Schiffman, J., Ma, X., et al. (2013). Tobacco smoke exposure and the risk of childhood acute lymphoblastic and myeloid leukemias by cytogenetic subtype. *Cancer Epidemiology, Biomarkers, & Prevention, 22*(9), 1600–1611. doi:10.1158/1055-9965.EPI-13-0350.

Michopoulos, F., Theodoridis, G., Smith, C. J., & Wilson, I. D. (2010). Metabolite profiles from dried biofluid spots for metabonomic studies using UPLC combined with oaToF-MS. *Journal of Proteome Research, 9*(6), 3328–3334. doi:10.1021/pr100124b.

Mitchell, F., & Shackleton, C. H. (1969). The investigation of steroid metabolism in early infancy. In O. Bodansky & C. P. Stewart (Eds.), *Advances in clinical chemsitry* (Vol. 12, pp. 141–215). New York: Academic Press.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. doi:10.1111/j.2041-210x.2012.00261.x.

Niinivirta, K., Isolauri, E., Laakso, P., Linderborg, K., & Laitinen, K. (2011). Dietary counseling to improve fat quality during pregnancy alters maternal fat intake and infant essential fatty acid status. *The Journal of Nutrition, 141*(7), 1281–1285. doi:10.3945/jn.110.137083.

Nochera, C., Goossen, L., Brutus, A., Cristales, M., & Eastman, B. (2011). Consumption of DHA + ePA by low-income women during pregnancy and lactation. *Nutrition in Clinical Practice: Official Publication of the American Society for Parenteral and Enteral Nutrition, 26*(4), 445–450.

Patti, G. J., Tautenhahn, R., & Siuzdak, G. (2012). Meta-analysis of untargeted metabolomic data from multiple profiling experiments. *Nature Protocols, 7*(3), 508–516. doi:10.1038/nprot.2011.454.

Prentice, P., Koulman, A., Matthews, L., Acerini, C. L., Ong, K. K., & Dunger, D. B. (2015). Lipidomic analyses, breast- and formula-feeding, and growth in infants. *Journal of Pediatrics, 166*(2), 276–281.e6. doi:10.1016/j.jpeds.2014.10.021.

Pupillo, D., Simonato, M., Cogo, P. E., Lapillonne, A., & Carnielli, V. P. (2016). Short-term stability of whole blood polyunsaturated fatty acid content on filter paper during storage at −28 °C. *Lipids, 51*(2), 193–198. doi:10.1007/s11745-015-4111-z.

R Core Team (2015). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/.

Rajesh, G. D., Bhat, B. V., Sridhar, M. G., & Ranganathan, P. (2000). Growth hormone levels in relation to birth weight and gestational age. *Indian Journal of Pediatrics, 67*(3), 175–177. doi:10.1007/BF02723656.

Raju, K. S. R., Taneja, I., Rashid, M., Sonkar, A. K., Wahajuddin, M., & Singh, S. P. (2016). DBS-platform for biomonitoring and toxicokinetics of toxicants: Proof of concept using LC-MS/MS analysis of fipronil and its metabolites in blood. *Scientific Reports, 6*, 22447. doi:10.1038/srep22447.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research, 43*(7), e47. doi:10.1093/nar/gkv007.

Sánchez-Guijo, A., Oji, V., Hartmann, M. F., Traupe, H., & Wudy, S. A. (2015). Simultaneous quantification of cholesterol sulfate, androgen sulfates, and progestagen sulfates in human serum by LC-MS/MS. *Journal of Lipid Research, 56*(9), 1843–1851. doi:10.1194/jlr.D061499.

Schindler, A., & Siiteri, P. (1968). Isolation and quantitation of steroids from normal human amniotic fluid. *Journal of Clinical Endocrinology and Metabolism Endocrinology, 28*, 1189098. doi:10.1210/jcem-28-8-1189.

Schwarz, E., Liu, A., Randall, H., Haslip, C., Keune, F., Murray, M., et al. (2009). Use of steroid profiling by UPLC-MS/MS as a second tier test in newborn screening for congenital adrenal hyperplasia: The Utah experience. *Pediatric Research, 66*(2), 230–235. doi:10.1203/PDR.0b013e3181aa3777.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscoape: A software environment for integrated models of biomolecular interaction networks. *Genome Research, 12*, 2498–2504. doi:10.1101/gr.1239303.

Smith, C., Elizabeth, J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry, 78*(3), 779–787. doi:10.1021/ac051437y.

Smith, C. A., O'maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., et al. (2005). METLIN A metabolite mass spectral database. *Proceedings of the 9th International Congress of Therapeutic Drug Monitoring & Clinical Toxicology, 27*(6), 747–751. doi:10.1097/01.ftd.0000179845.53213.39.

Tang, W. H. W., Wang, Z. E., Levison, B. S., Koeth, R. A., Britt, E. B., Fu, X. M., et al. (2013). Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *New England Journal of Medicine, 368*(17), 1575–1584. doi:10.1056/NEJMoa1109400.

Vu, D. H., Koster, R. A., Alffenaar, J. W. C., Brouwers, J. R. B. J., & Uges, D. R. A. (2011). Determination of moxifloxacin in dried blood spots using LC-MS/MS and the impact of the hematocrit and blood volume. *Journal of Chromatography B, 879*(15–16), 1063–1070. doi:10.1016/j.jchromb.2011.03.017.

Wang, Z., Klipfell, E., Bennett, B. J., Koeth, R., Levison, B. S., Dugar, B., et al. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature, 472*(7341), 57–63. doi:10.1038/nature09922.

Wiemels, J., Cazzaniga, G., Daniotti, M., Eden, O., Addison, G., Masera, G., et al. (1999). Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet, 354*(9189), 1499–1503. doi:10.1016/S0140-6736(99)09403-9.

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y., et al. (2013). HMDB 3.0-The Human Metabolome Database in 2013. *Nucleic Acids Research, 41*(D1), 801–807. doi:10.1093/nar/gks1065.

Youhnovski, N., Bergeron, A., Furtado, M., & Garofolo, F. (2011). Pre-cut dried blood spot (PCDBS): an alternative to dried blood spot (DBS) technique to overcome hematocrit impact. *Rapid Communications in Mass Spectrometry: RCM, 25*(19), 2951–2958. doi:10.1002/rcm.5182.