Cord Blood Metabolic Signatures of Birth Weight: A Population-Based Study

Peer-reviewed author version

Robinson, Oliver; Keski-Rahkonen, Pekka; Chatzi, Leda; Kogevinas, Manolis; NAWROT, Tim; Pizzi, Costanza; PLUSQUIN, Michelle; Richiardi, Lorenzo; Robinot, Nivonirina; Sunyer, Jordi; Vermeulen, Roel; Vrijheid, Martine; Vineis, Paolo; Scalbert, Augustin & Chadeau-Hyam, Marc (2018) Cord Blood Metabolic Signatures of Birth Weight: A Population-Based Study. In: JOURNAL OF PROTEOME RESEARCH, 17 (3), p. 1235-1247.

DOI: 10.1021/acs.jproteome.7b00846
Handle: http://hdl.handle.net/1942/26605

Journal of proteome research

Article

Cord blood metabolic signatures of birthweight: a population based study

Oliver Robinson, Pekka Keski-Rahkonen, Leda Chatzi, Manolis Kogevinas, Tim Nawrot, Costanza Pizzi, Michelle Plusquin, Lorenzo Richiardi, Nivonirina Robinot, Jordi Sunyer, Roel Vermeulen, Martine Vrijheid, Paolo Vineis, Augustin Scalbert, and Marc Chadeau-Hyam

J. Proteome Res., **Just Accepted Manuscript** • DOI: 10.1021/acs.jproteome.7b00846 • Publication Date (Web): 05 Feb 2018

ACS Publications

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

Cord blood metabolic signatures of birthweight: a population based study

Oliver Robinson[a]‡, Pekka Keski-Rahkonen[b]‡, Leda Chatzi[c-e], Manolis Kogevinas[f-h], Tim Nawrot[i-j], Costanza Pizzi[k], Michelle Plusquin[i-j], Lorenzo Richiardi[k], Nivonirina Robinot[b], Jordi Sunyer[f-h], Roel Vermeulen[l], Martine Vrijheid[f-h], Paolo Vineis[a], Augustin Scalbert[b], Marc Chadeau-Hyam[a]*

a. MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK

b. International Agency for Research on Cancer (IARC), Lyon, France

c. Department of Social Medicine, Faculty of Medicine University of Crete, Heraklion, Greece.

d. Department of Preventive Medicine, Keck School of Medicine, University of South California, Los Angeles, CA

e. Department of Genetics & Cell Biology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

f. ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain

g. Universitat Pompeu Fabra (UPF), Barcelona, Spain

h. CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

i. Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium

j. Department of Public Health & Primary Care, Leuven University, Leuven, Belgium.

k. Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin and CPO-Piemonte, Turin, Italy

l. Utrecht University, Institute for Risk Assessment Sciences, Environmental Epidemiology Division, Utrecht, Netherlands

KEYWORDS: Metabolomics; birthweight; fetal growth; cord blood; metabolism; pathway perturbation

ABSTRACT: Birthweight is an important indicator of maternal and fetal health, and a predictor of health in later life. However, the determinants of variance in birthweight are still poorly understood. We aimed to identify the biological pathways, which may be perturbed by environmental exposures, that are important in determining birthweight. We applied untargeted mass-spectrometry based metabolomics to 481 cord blood samples collected at delivery in four birth cohorts from across Europe: ENVIRONAGE (Belgium), INMA (Spain), Piccolipiu (Italy) and Rhea (Greece). We performed a metabolome-wide association scan for birthweight on over 4000 metabolic features, controlling the false discovery rate at 5%. Annotation of compounds was conducted through reference to authentic standards. We identified 68 metabolites significantly associated with birthweight, including vitamin A, progesterone, docosahexaenoic

acid, indolelactic acid, and multiple acylcarnitines and phosphatidylcholines. We observed enrichment ($p < 0.05$) of the tryptophan metabolism, prostaglandin formation, C21-steroid hormone signalling, carnitine shuttle and glycerophospholipid metabolism pathways. Vitamin A was associated with both maternal smoking and birthweight, suggesting a mediation pathway. Our findings shed new light on the pathways central to fetal growth and will have implications for antenatal and perinatal care and potentially for health in later life.

Introduction

Weight at birth is of importance due to its relation both to the health of the mother and the new born [1] and is a predictor of the subsequent development of the child [2 3]. Low birthweight has also been associated with disease in later life, including cardiovascular disease and diabetes[4] while high birthweight is associated with increased risk of developing breast cancer and other malignancies[5-6]. Thus, under the developmental origin of health and disease hypothesis, adverse fetal development may have a lifelong impact. Exposure to a number of different agents during pregnancy including smoking[7], air pollution[8] and chemicals such as polychlorinated biphenyls [9] have been associated with lower birthweight. Improved understanding of the biological pathways associated with environmentally induced alterations in birthweight may identify mechanisms through which fetal growth is affected. This may in-turn both inform primary care and ultimately improve the causal evidence regarding adverse fetal development.

Metabolomics is increasingly used in maternal-fetal medicine[10] to identify biological changes associated with fetal growth. Horgan *et al.*[11] examined ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS) profiles of first trimester maternal plasma samples to identify metabolites predictive of small for gestational age babies. Maitre *et al.* used nuclear magnetic resonance (NMR) spectroscopy on maternal pregnancy urine samples to identify metabolites predictive of preterm birth, small for gestational age, and fetal growth restriction[12] and birthweight[13]. Dessi *et al.*[14] identified four metabolites associated with fetal growth restriction in neonatal urine samples. However, only a handful of studies have investigated metabolic changes in cord blood, which is a particularly relevant tissue as it contains the essential nutrients, hormones and immunological factors, and potentially harmful xenobiotic metabolites, to which the developing fetus is directly exposed. Horgan et al.[11] compared six

small for gestational age babies with controls observing differences in levels of sphingolipids, phospholipids and carnitines. Ivorra *et al*.[15] and Tea *et al*.[16] both used nuclear magnetic resonance spectroscopy to compare a small number of low birthweight and very low birthweight newborns respectively with controls and detected some differences in metabolite levels. Recently, Hellmuth *et al*[17] applied a targeted mass spectrometry based analysis and observed a positive association between lysophosphatidylcholines (lysoPCs) and birthweight. While initial studies have so far been based on small samples or limited subsets of molecules, they demonstrate the potential of metabolic profiling to detect biological pathways related to fetal development. Adequately powered studies using sensitive and untargeted platforms are now required.

In this study, we have employed untargeted UHPLC-MS -based metabolomics to identify metabolic features associated with birthweight in cord blood collected from a large population-based sample from four European birth cohorts. We aimed to understand the mechanisms that are important to fetal growth, which may be influenced by the maternal environment including exposure to air pollutants and tobacco smoke, and that may impact on health over the life course.

Experimental Section

Cohorts and sampling

Within the context of the EXPOsOMICS collaborative European project[18], metabolomic analyses were conducted on umbilical cord blood samples from four population-based birth cohorts: ENVIR*ON*AGE[19], Piccolipiu[20], INMA[21] and Rhea[22]. The ENVIR*ON*AGE cohort recruited women when they arrived for delivery at the South-East-Limburg Hospital in Gent, Belgium between 2010 and 2013. The INMA cohort enrolled women during the first trimester of

pregnancy at public primary health care centres or hospitals in Sabadell, Spain between 2004 and 2006. The participating Piccolipiu cohort centre enrolled women giving birth at the main hospital of the City of Turin, Italy, between 2011 and 2013. The Rhea cohort enrolled women during the first trimester of pregnancy at public primary health care centres or hospitals in Heraklion, Greece between 2007 and 2008. Whole blood samples were collected using venipuncture of cord vessels before the placenta was delivered and processed as follows in each cohort: In ENVIR*ON*AGE, samples were collected into EDTA (BD, Franklin Lakes, NJ, USA) vacutainers and within 20 mins were centrifuged at 3,200 rpm for 15 min into plasma; In Piccolipiu, samples were collected into BD EDTA vacutainers, stored at 4°C for less than 24 h and centrifuged for 10 minutes at 1300g into plasma; In Rhea samples were collected into BD gel separator vacutainers and centrifuged within 2 h at 2500 rpm for 10 min into serum; In INMA samples were collected into BD gel separator vacutainers, stored at 4°C for less than 4 h and centrifuged at 3000 rpm for 10-15 min into serum. Samples were immediately frozen at -20 °C (INMA) or -80 °C (all other cohorts) until analysis. Cohort inclusion criteria and further protocols can be found in the respective cohort references.

Family lifestyle factors were collected from mothers through interview by trained fieldworkers and medical history for each family transferred from hospital records. Samples were selected from each cohort on the basis of biomaterial and data availability. Selected samples were shipped to International Agency for Research on Cancer, Lyon, France for metabolomics analysis.

Exposure Assessment

Exposure to air pollutants (particulate matter ≤ 10 microns (PM_{10}), ≤ 2.5 microns ($PM_{2.5}$) and ≤ 0.1 microns (UFP) and NO_2) was assessed at the home address, averaged over the year before pregnancy, by land use regression models[23].

Sample analysis

Samples were randomized and prepared by mixing a 30-µL aliquot with 200 µL of acetonitrile and filtering the precipitate with 0.2 µm Captiva ND plates (Agilent Technologies). The filtrate was collected into a polypropylene well plate that was sealed and kept refrigerated until analysis. A quality control (QC) sample was prepared by mixing small aliquots of 91 randomly selected study samples. Aliquots (30 µL) of the QC sample were then processed along with the study samples, with each 96-well plate containing four individually extracted QCs. Samples were analyzed as a single uninterrupted batch with a UHPLC-MS system consisting of a 1290 Binary LC system, a Jet Stream electrospray ionization (ESI) source, and a 6550 QTOF mass spectrometer (Agilent Technologies). Autosampler tray was kept refrigerated and 2 µL of the sample solution was injected on an ACQUITY UPLC HSS T3 column (2.1 × 100mm, 1.8 µm; Waters). Column temperature was 45 °C and mobile phase flow rate 0.4 ml/min, consisting of ultrapure water and LC-MS grade methanol, both containing 0.05 % (v/v) of formic acid. The gradient profile was as follows: 0–6 min: 5% → 100% methanol, 6–10.5 min: 100% methanol, 10.5–13 min: 5% methanol. The mass spectrometer was operated in positive polarity using the following conditions: drying gas (nitrogen) temperature 175 °C and flow 12 L/min, sheath gas temperature 350 °C and flow 11 L/min, nebulizer pressure 45 psi, capillary voltage 3500 V, nozzle voltage 300 V, and fragmentor voltage 175 V. Data was acquired in centroid format using an extended dynamic range mode, with a scan rate of 1.67 Hz and a mass range from 50 to 1000. For MS/MS analyses the isolation width was 1.3 Da and collision energies 10, 20, and 40 V. Continuous mass axis calibration was performed using two reference ions (*m/z* 121.050873 and *m/z* 922.009798). The analytical run was initiated with priming injections of in-house human plasma extract to achieve a stable response, followed by the study samples and one QC sample after every 12 injections. Data was acquired using MassHunter Acquisition B.05.01 software.

Data pre-processing

Pre-processing of the acquired data was performed using Qualitative Analysis B.06.00 SP1, DA Reprocessor, and Mass Profiler Professional 12.1 software (Agilent Technologies). The initial processing was performed with Molecular Feature Extraction (MFE) algorithm for small molecules using a mass range of 50-1000. Thresholds for the mass and chromatographic peak heights were 1500 and 10000 counts, respectively. Quality score threshold was 80. Only singly charged proton adducts ($[M+H]^+$) were included. Spacing tolerance for isotope peaks was 0.0025 m/z plus 7 ppm. Isotope model for common organic molecules was used and features with indeterminate neutral mass were excluded. Feature alignment was performed with retention time and mass windows of 0.075 min and 15 ppm + 2 mDa. A target list for a recursive extraction was created from features present in at least 2% of the samples. Find by Formula (FBF) algorithm was then employed with match tolerance for the mass and retention time ±10 ppm and ±0.04 min, respectively. Ion species was limited to $[M+H]^+$ and a threshold for chromatographic peak height was 2000 counts. The resulting features were aligned in Agilent Mass Profiler Professional using the same parameters as described above. For statistical analysis, metabolic features present in less than 60% of the samples were removed, data were log-transformed and missing values were imputed using the impute.QRILC function within the imputeLCMD R package[24].

Statistical analysis

The relationship between birthweight and the cord metabolome was assessed using a metabolome-wide association scan (MWAS) approach with separate linear regression models for each metabolic feature using the 'omics' R package[25]. To account for multiple comparisons, we applied a Benjamini- Hochberg correction using an overall False Discovery Rate of (FDR) below 5%. The covariates included in the main MWAS analysis were gestational age, sex, cohort,

maternal height, maternal weight, and paternal height. Covariates were chosen first following visualisation of assumptions using a directed acyclical graph (figure s1) and then following testing of associations with both birthweight in bivariate analyses (Analysis of Variance or Pearson's correlation tests) and with metabolic features as visualised by Q-Q plots of the p-value distribution. We chose to not adjust for environmental factors that may be associated with birthweight in the main analysis since we hypothesised that the metabolome may mediate these associations. Instead we tested for potential confounding by environmental factors in sensitivity analyses through stratification by covariates and adjustment separately for each birthweight associated risk factor.

We further investigated links between risk factors and birthweight by first constructing a linear regression model of birthweight, including those factors associated with birthweight ($p \leq 0.1$) in bivariate analyses. We then adjusted the model for the metabolome, using the first components of a principal component analysis (PCA) of metabolites associated with birthweight in the main MWAS analysis. The number of components to include was selected by examination of a scree plot of explained variance. Where risk factor associations were attenuated following adjustment on birthweight-related metabolites, we tested the association of that factor with all birthweight related metabolites in a further analysis.

All analyses were performed in R version 3.3 [26].

Metabolite and pathway annotation

Annotation of the discriminating features was done as follows: 1) The *m/z* values of all the features were searched against the human metabolite database[27]and Metlin[28], using $[M+H]^+$, $[M+Na]^+$ and $[M+2H]^{2+}$ as adducts and ±8ppm for molecular weight tolerance. In addition, MyCompoundID metabolite library[29] was searched for potential conjugates (sulphates, glucuronides) and neutral losses ($-NH_3$, H_2O), using $[M+H]^+$ ions and ±8 ppm mass tolerance. 2)

Features were grouped based on retention time similarity and intensity correlation across the samples to assist in identifying ions originating from the same metabolite. 3) Quality of the chromatographic peaks and spectra were inspected and the plausibility of database candidates was assessed based on retention time, isotope pattern and adduct formation or neutral losses. 4) Identification was confirmed by reanalysis of representative samples and pure standards when available, and comparison of the retention times and the MS/MS spectra. When standards were not available, MS/MS spectra were acquired and compared against those in mzCloud (www.mzcloud.org) or Metlin. Chromatograms and spectra can be found in the supporting information (Additional dataset 1). The level of identification was as proposed by Sumner et al [30].

Significantly enriched metabolic pathways were identified using the *mummichog* program[31]. The algorithm searches tentative compound lists from metabolite reference databases against an integrated model of human metabolism to identify functional activity. Fisher's exact tests and permutation are used to infer p-values for likelihood of pathway enrichment among significant features as compared to pathways identified among the entire compound set present in reference list (the entire metabolome dataset), considering the probability of mapping the significant m/z features to pathways. *Mummichog* parameters were set to match against ions produced by the MS method employed ($[M]^+$, $[M+H]^+$, $[M+2H]^{2+}$, $[M+3H]^{3+}$, [M+Na]+, $[M+H+Na]^{2+}$, $[M+K]^+$, $[M-H_2O+H]^+$, $[M-H_4O_2+H]^+$, $[M-NH_3+H]^+$, $[M-CO+H]^+$, $[M- CO_2+H]^+$, $[M-HCOOH+H]^+$, $[M+HCOONa]^+$, $[M-HCOONa+H]^+$, $[M-C_3H_4O_2+H]^+$) at ± 8 ppm mass tolerance.

Results

Participant information

Samples of cord blood collected from 499 deliveries (200 from ENVIR*ON*AGE, 100 each from INMA and Rhea, and 99 from Piccolipiu) were included in the analysis. The mean

birthweight was 3309 g (interquartile range 2992 – 3598 g) with 16 (3 %) babies born with low birthweight (< 2500g). Demographic information and covariate associations with birthweight are shown in table 1. Gestational age, cohort, sex, maternal height, weight and body mass index, paternal height, maternal weight gain during pregnancy, smoking by the mother and residential exposure to air pollution were all significantly associated with birthweight in bivariate analyses.

Table 1. Participant information and covariate associations with birthweight.

	N (%) or mean (IQR)	Missing values	mean birthweight (IQR) or *r* with birthweight	p value*
Birthweight (g)	3309 (2992 - 3598)	1	-	-
Gestational age (wks)	39.2 (38.1 -40)	1	0.21	<0.0001
Cohort		0		0.014
Rhea	100 (20)	-	3220 (2970 - 3520)	
Environage	200 (40)	-	3384 (3045 - 3705)	
Piccolipiu	99 (20)	-	3221 (2955 - 3490)	
Inma	100 (20)	-	3298 (2998 - 3572)	
Sex		1		0.0005
Male	258 (52)	-	3377 (3076 -3648)	
Female	240 (48)	-	3237 (2915 - 3550)	
Parity		4		0.13
0	129 (26)	-	3275 (2990 - 3570)	
1	238 (48)	-	3294 (2971 - 3579)	
2	128 (260	-	3379 (3119 - 3640)	
Season of conception		3		0.35
January-March	122 (25)	-	3318 (3000 - 3550)	
April-June	129 (26)	-	3296 (2970 - 3600)	
July-September	160 (32)	-	3348 (3019 - 3630)	
October-December	85 (17)	-	3240 (2980 - 3485)	
Delivery		2		0.18
Vaginal	382 (77)	-	3324 (3000 - 3614)	

Caesarean	115 (23)	-	3259 (2960 – 3542)	
Maternal age (years)	30.7 (27.9-34)		-0.04	0.44
Maternal height (cm)	164.5 (160-168)	2	0.27	<0.0001
Maternal weight (kg)	62 (42 - 130)	2	0.24	<0.0001
Maternal BMI	23 (20.8 - 25.8)	2	0.15	0.001
Maternal weight gain (kg)	13.63 (10-17)	13	0.21	0.0002
Paternal height (cm)	177.8 (173 -182)	17	0.14	0.003
Paternal weight (kg)	82 (73 - 90)	17	0.07	0.11
Mother born in country		9		0.69
Yes	432 (88)	-	3315 (3000 - 3586)	
No	58 (12)	-	3287 (2881 - 3638)	
Mother's education		12		0.64
Primary school	62 (13)	-	3291 (2975 - 3595)	
Secondary School	209 (43)	-	3296 (2980 - 3590)	
University of higher	216 (44)	-	3333 (3060 - 3589)	
Father's Education		28		0.08
Primary school	96 (20)	-	3236 (2915 - 3550)	
Secondary School	227 (47)	-	3288 (2982 - 3580)	
University of higher	148 (30)	-	3365 (3084 - 3600)	
Maternal Smoking (during 2nd trimester)		3		0.06
Yes	57 (11%)	-	3207 (2890 - 3440)	
No	439 (89%)	-	3324 (3000 - 3612)	
Passive smoke exposure		15		0.50
Yes	172 (36)	-	3295 (2970 - 3616)	

No	312 (64)	-	3324 (3000 - 3592)	
Exposure to PM_{10} ($\mu g/m^3$)	33.0 (17.8 - 44.0)	4	-0.13	0.002
Exposure to $PM_{2.5}$($\mu g/m^3$)	18.14 (12.4 -21.0)	4	-0.11	0.02
Exposure to NO_2 ($\mu g/m^3$)	29.1 (13.4 – 40.3)	4	-0.11	0.02

*p value for association with birthweight, calculated from Pearson's correlation test (continuous variable) or analysis of variance test (categorical variable). IQR = Interquartile range.

Metabolomic data

The total number of metabolomic features was 9947. Out of these, 4870 (49%) were found in at least 90% of the QC sample injections, with 4019 (83%) having a coefficient of variation below 30%, indicating good reproducibility. Features were excluded if present in less than 60% of participant samples, leaving 4714 metabolic features for statistical analysis. A principal component analysis (PCA) of these features (whole metabolome PCA) showed that the 1st component explained 18% of variance and 307 components explained 95% of the variance suggesting considerable redundancy within the data (figure s2A). There was clear separation along the second component by cohort (figure s2B). No separation was observed with other covariates or technical variables.

Metabolic features associated with birthweight

In the main MWAS analysis of 481 participants with complete covariate data, adjusted for gestational age, cohort, sex, maternal height, maternal weight, and paternal height, 138 features were significantly associated (FDR < 5%) with birthweight (figure 1).

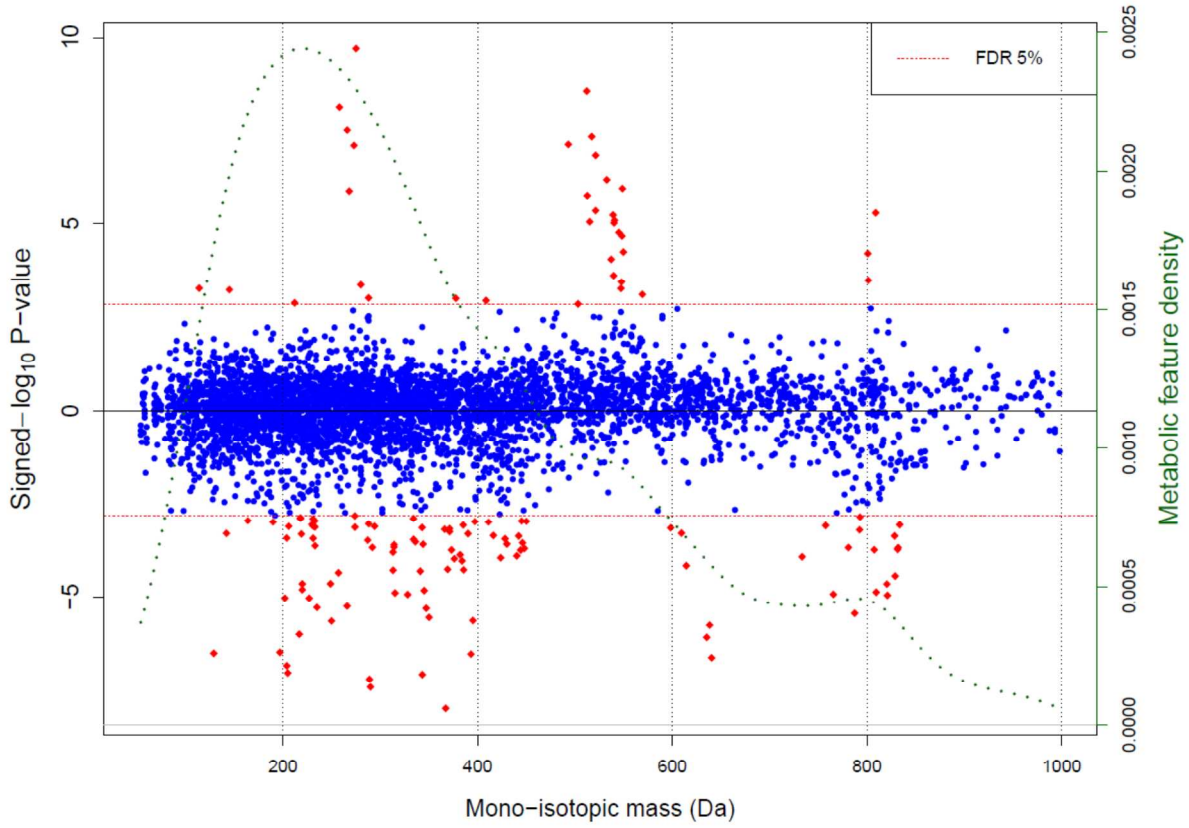


Figure 1: Manhattan Plot showing strength of association (log_{10} p value) with birthweight for each metabolic feature, arranged by mono-isotopic mass. Green dotted line shows density of metabolic features. Sign of p value indicates direction of association. Features are coloured red if they are metabolome-wide significant at the 5% false discovery rate.

After grouping of ions originating from the same molecule (matched by retention time and pairwise feature correlation), there were 68 unique compounds associated with birthweight. 29 of these were annotated to at least the level of compound class, based on matching mass, isotope pattern, and product ion spectra (table 2). Fifteen compounds were identified as acylcarnitines, nine as phosphatidylcholines (PCs) or lysoPCs, two as tryptophan metabolites (indolelactic acid and an isomeric form of methoxykynurenate), two as essential nutrients (retinol (vitamin A) and

docosahexaenoic acid (DHA)) and one as a steroid hormone. Seven of these were conclusively identified through comparisons with authentic standards: decanoylcarnitine, dodecanoylcarnitine and tetradecanoylcarnitine, indolelactic acid, retinol, the ω-3 fatty acid DHA and progesterone, one of the major progestational steroid hormones. In addition to the 29 annotated metabolites, five compounds (three diacylglycerols and two lysoPCs) were tentatively assigned based on exact mass and isotope pattern, due to low intensity or lack of MS/MS spectra or authentic standards (table 2). Retention time and exact mass of all significantly associated features, including unassigned compounds, is given in table s1. Chromatograms and mass spectra of all annotated compounds are given in additional dataset 1.

Table 2: Metabolites significantly associated with birthweight

m/z	Retention time (minutes)	Ion	Name[a]	ID level[b]	Compound class	Direction of change with increasing birthweight	P value
368.2793	5.631	$[M+H]^+$	3, 5-Tetradecadiencarnitine (C14:2)	2	Acylcarnitine	DOWN	1.00E-08
518.3216	6.782	$[M+H]^+$	LysoPC (C18:3)	2	Glycerophospholipid	UP	4.63E-08
494.3250	6.817	$[M+H]^+$	LysoPC (C16:1)	2	Glycerophospholipid	UP	7.45E-08
344.2797	5.647	$[M+H]^+$	Dodecanoylcarnitine (C12:0)	1	Acylcarnitine	DOWN	8.31E-08
206.0822	3.829	$[M+H]^+$	Indolelactic acid	1	Tryptophan metabolite	DOWN	9.16E-08
522.3555	6.980	$[M+H]^+$	LysoPC (C18:1)	2	Glycerophospholipid	UP	1.45E-07
641.5112	9.938	$[M+Na]^+$	Diacylglycerol (C36:3)	4	Glycerolipid	DOWN	2.34E-07
269.2278	7.219	$[M-H_2O+H]^+$	Retinol	1	Vitamin	UP	1.30E-06
639.4946	9.408	$[M+Na]^+$	Diacylglycerol (C36:4)	4	Glycerolipid	DOWN	1.74E-06
396.3100	5.944	$[M+H]^+$	9,12-Hexadecadienoylcarnitine (C14:1)	2	Acylcarnitine	DOWN	2.32E-06
522.3565	7.058	$[M+H]^+$	LysoPC (C18:1)	2	Glycerophospholipid	UP	4.55E-06
329.2482	7.232	$[M+H]^+$	Docosahexaenoic acid	1	Fatty Acid	DOWN	1.18E-05
766.5815	8.859	$[M+H]^+$	PC-O (C36:4) ($C_{44}H_{80}NO_7P$)	3	Glycerophospholipid	DOWN	1.20E-05

316.2489	5.139	$[M+H]^+$	Decanoylcarnitine (C10:0)	1	Acylcarnitine	DOWN	1.30E-05
810.6053	9.169	$[M+H]^+$	PC (C38:4)	3	Glycerophospholipid	DOWN	1.37E-05
258.1699	2.831	$[M+H]^+$	2-Hexenoylcarnitine (C6:1)	2	Acylcarnitine	DOWN	4.43E-05
342.2641	5.422	$[M+H]^+$	Trans-2-Dodecenoylcarnitine (C12:1)	2	Acylcarnitine	DOWN	4.89E-05
314.2318	4.967	$[M+H]^+$	Decenoylcarnitine_2[c] (C10:1)	2	Acylcarnitine	DOWN	5.20E-05
370.2955	5.840	$[M+H]^+$	Tetradecenoylcarnitine (C14:1)	2	Acylcarnitine	DOWN	5.49E-05
615.4959	9.763	$[M+Na]^+$	Diacylglycerol (C34:2)	4	Glycerolipid	DOWN	6.82E-05
734.5700	8.960	$[M+H]^+$	PC (C32:0)	3	Glycerophospholipid	DOWN	1.17E-04
314.2321	4.878	$[M+H]^+$	Decenoylcarnitine_1[c] (C10:1)	2	Acylcarnitine	DOWN	1.55E-04
315.2320	6.394	$[M+H]^+$	Progesterone	1	Steroid hormone	DOWN	2.11E-04
288.2171	4.422	$[M+H]^+$	L-Octanoylcarnitine (C8:0)	2	Acylcarnitine	DOWN	3.21E-04
232.1537	1.927	$[M+H]^+$	Butyrylcarnitine / Isobutyryl-L-carnitine (C4:0)	2	Acylcarnitine	DOWN	3.61E-04
220.0605	3.671	$[M+H]^+$	Methoxykynurenate[d] ($C_{11}H_9NO_4$)	3	Tryptophan metabolite	DOWN	4.69E-04
548.3681	7.141	$[M+H]^+$	LysoPC (C20:2)	4	Glycerophospholipid	UP	5.58E-04
782.5712	8.628	$[M+H]^+$	PC (C36:4)	3	Glycerophospholipid	DOWN	6.03E-04
372.3112	6.006	$[M+H]^+$	Tetradecanoylcarnitine (C14:0)	1	Acylcarnitine	DOWN	6.46E-04
758.5747	8.684	$[M+H]^+$	PC (C34:2)	3	Glycerophospholipid	DOWN	7.88E-04
570.3551	7.021	$[M+H]^+$	LysoPC (C22:5)	4	Glycerophospholipid	UP	8.01E-04

386.2899	5.568	$[M+H]^+$	3-Hydroxy-cis-5-tetradecenoylcarnitine (C14:1)	2	Acylcarnitine	DOWN	8.10E-04
412.3045	5.750	$[M+H]^+$	3-Hydroxyhexadecadienoylcarnitine (C16:1)	2	Acylcarnitine	DOWN	9.82E-04
398.3264	6.109	$[M+H]^+$	Hexadecenoylcarnitine (C16:1)	2	Acylcarnitine	DOWN	1.01E-03

[a] Common names as used in HMDB. Molecular formulas given when identification inconclusive or more than one isomers are possible. Chain lengths and number of double bonds are indicated for acylcarnitines, glycerophospholipid, and diacylglycerols.

[b] Level 1 (identified compounds): retention time and MS/MS spectra matches with an authentic chemical standard; Level 2 (putatively annotated compound): no standard available or analysed but has a single database candidate within 5 ppm mass error, matching isotope pattern and MS/MS spectra; Level 3 (putatively characterized compound class): MS/MS spectral similarity with compounds from a known chemical class; Level 4 (Unknown compounds): no standard or MS/MS spectra available but a chemically plausible hit in a metabolite database within 5 ppm mass error and a matching isotope pattern[60]

[c] Two isomers of decenoylcarnitine were identifed.

[d] Likely a positional isomer of methoxykynurenate based on closeness of retention time and similarities in spectra to those acquired from authentic standard of 8-methoxykynurenate.

As shown in clustered correlation heatmap (figure 2), the longer chained acylcarnitines clustered together and were moderately correlated with progesterone, the tryptophan metabolites and the shorter chained acylcarnitines. The PCs, smaller sized lysoPCs and tentatively assigned diacylglycerols formed three further clusters. The diacylglycerol cluster was moderately correlated with DHA. Both the diacylglycerol cluster and DHA were negatively correlated with the cluster of smaller sized lysoPCs.

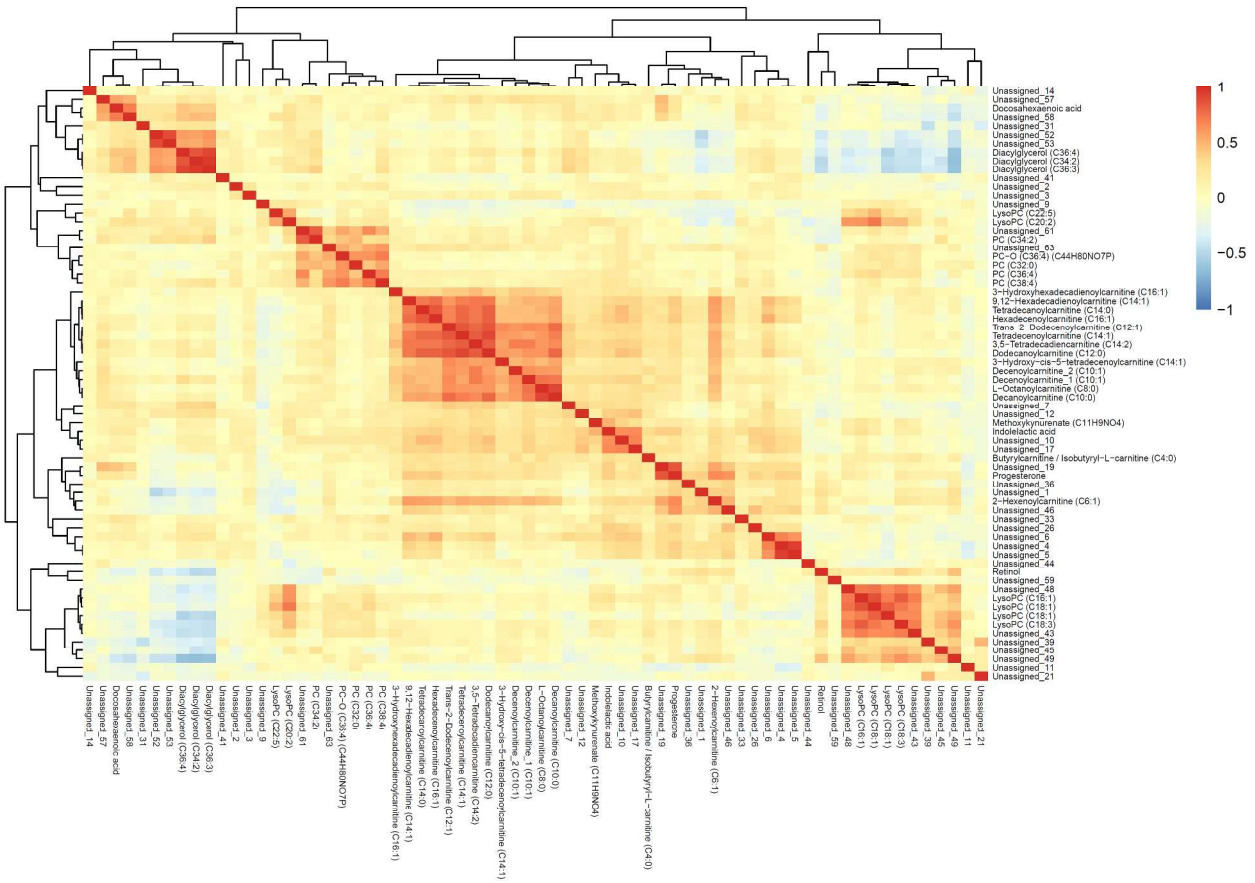


Figure 2: Clustered correlation heat map of metabolites significantly (false discovery rate < 5%) related to birthweight. Numbers following unassigned compounds refer to compound numbers (see table s1).

In PCA of the significant compounds, over 50% of variance was explained by 5 components, and 41 components were required to explain over 95% of the variance of these features. There was some separation in results by cohort (figure 3a) and inspection of the component loadings (figure 3b) revealed this was driven mainly by differing levels of acylcarnitines along the first component and of lysoPCs and the tentatively assigned diacylglycerols along the second component.

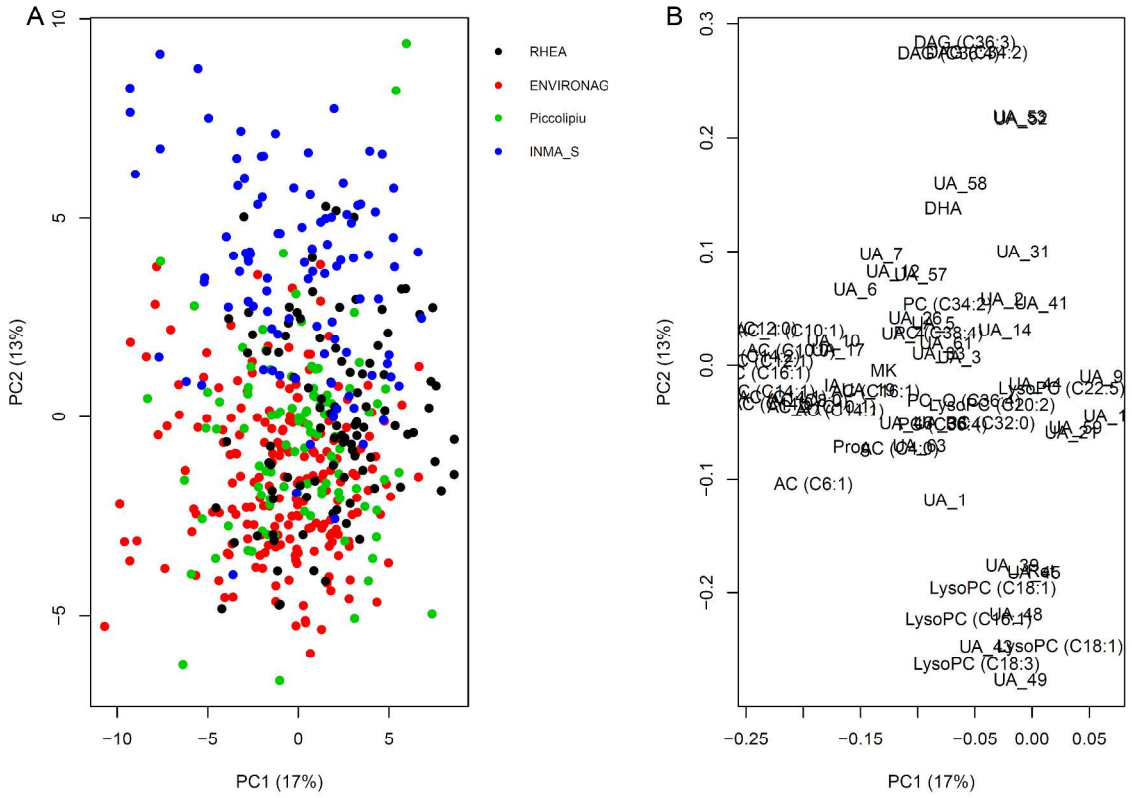


Figure 3: PCA analysis of significant metabolites. A: Scores scatter plot of first two components, coloured by cohort. B: Loadings scatterplot of first two components. Abbreviations: AC: acylcarnitine, UA: Unassigned, PC: Phosphatidylcholine, LysoPC: lysophosphatidylcholine, Prog: Progesterone, Ret: Retinol, DAG: Diacylglycerol, IA: Indolelactic acid, MK:

methoxykynurenate, DHA: Docosahexaenoic acid. Numbers following unassigned compounds refer to compound numbers (see table s1).

Sensitivity analysis

To further assess the potential cohort-specific results, we also ran our models on each cohort separately (table s2). Additionally, we stratified by sex and re-ran analyses after excluding preterm and caesarean deliveries table s2). Results were generally consistent across cohorts, although we noted stronger associations for the smaller sized lysoPCs in the INMA cohort and non-significant and opposite directions of association for some acylcarnitines (dodecanoylcarnitine, decanoylcarnitine, Trans-2-dodecenoylcarnitine, and tetradecenoylcarnitine) in Rhea and for the three tentatively assigned diacylglycerols in Piccolipiu. These differences largely reflect the heterogeneity of the metabolic profiles we observed in each cohort (figure 3). Upon stratification by sex, we observed stronger associations for decanoylcarnitine, tetradecanoylcarnitine and hexadecenoylcarnitine in girls and a stronger association with PC ($C_{40}H_{80}NO_8P$) in boys. Associations were similar when preterm and caesarean deliveries were excluded, except for slightly stronger associations with decanoylcarnitine and dodecanoylcarnitine in analysis of vaginal deliveries only.

To test for the potential confounding of our results we additionally adjusted the main model separately by factors that influence birthweight. Additional factors included exposure to PM_{10}, paternal education, smoking continuing into the second trimester and maternal weight gain. We also checked for the confounding effects of mode of delivery and cohort differences in the metabolome data, using the second component of the whole metabolome PCA. As shown in table s3, associations were robust to adjustment by these factors, indicating negligible confounding.

To investigate the role of the metabolome in potentially mediating the effect of risk factors on birthweight, we adjusted a risk factor model of birthweight (including exposure to PM_{10}, maternal and paternal education, smoking continuing into the second trimester and maternal weight gain) by the first five principal components of birthweight associated metabolites (explaining 50.2% of the total variance in the set of 68 metabolites). Of the included risk factors, only the effect size estimate linking birthweight and smoking continuing into the second trimester was decreased upon adjustment on birthweight-related metabolites (table 3). Specifically, while the unadjusted model estimated that babies of mothers who smoked during the second trimester weighed 88.5g (95% confidence interval (CI): -197.6g, 20.5g) less at birth than babies of mothers who did not, that weight loss dropped to 74.0g (95% CI: -170.7g, 22.6g) in the model adjusting for birthweight-related metabolites.

Table 3. Associations of risk factors with birthweight, with and without adjustment for metabolites

	Risk factor model[a]	**Risk factor model, adjusted for birthweight associated metabolome[b]**	**Risk factor model, adjusted for retinol levels in cord blood[c]**
	β (95% CI)	β (95% CI)	β (95% CI)
Exposure to PM_{10} during pregnancy	-3.37 (-11.74, 5.00)	-4.97 (-12.40, 2.45)	-2.90 (-11.09, 5.28)
Education level of father (ref. primary school)			
Secondary school	59.7 (-29.56, 148.96)	80.83 (2.11, 159.54)	71.77 (-15.65, 159.19)
University or higher	116.08 (16.18, 215.98)	126.04 (38.11, 213.97)	119.19 (21.52, 216.86)

Mother smoked during second trimester	-88.54 (-197.62, 20.54)	-74.02 (-170.69, 22.65)	-45.06 (-153.30, 63.18)
maternal weight gain	12.42 (5.77, 19.07)	13.87 (8.00, 19.75)	11.69 (5.18, 18.19)

a. Adjusted for gestational age, cohort, maternal height and pre-pregnancy weight, sex of baby and paternal height

b. Adjusted for gestational age, cohort, maternal height and pre-pregnancy weight, sex of baby and paternal height and first 5 principal components of PCA on all birthweight-associated metabolites.

c. Adjusted for gestational age, cohort, maternal height and pre-pregnancy weight, sex of baby and paternal height and levels of retinol in cord blood

Linear models investigating the association between the birthweight-associated metabolites and maternal smoking, adjusted for cohort, maternal education and body mass index, showed that only blood retinol levels were significantly different (lower, $p = 4.26 \times 10^{-4}$) among mothers who smoked during the second trimester (figure 4). Based on this observation, we additionally looked at the evolution of association between birthweight risk factors in a model adjusting for retinol levels (table 3). The effect size estimate of smoking on birthweight was halved upon adjustment on retinol levels (β= -45.1g, 95% CI: -153.3g, 63.2g). Although these changes are not statistically significant, they indicate that retinol may partly mediate the effects of smoking on birthweight.

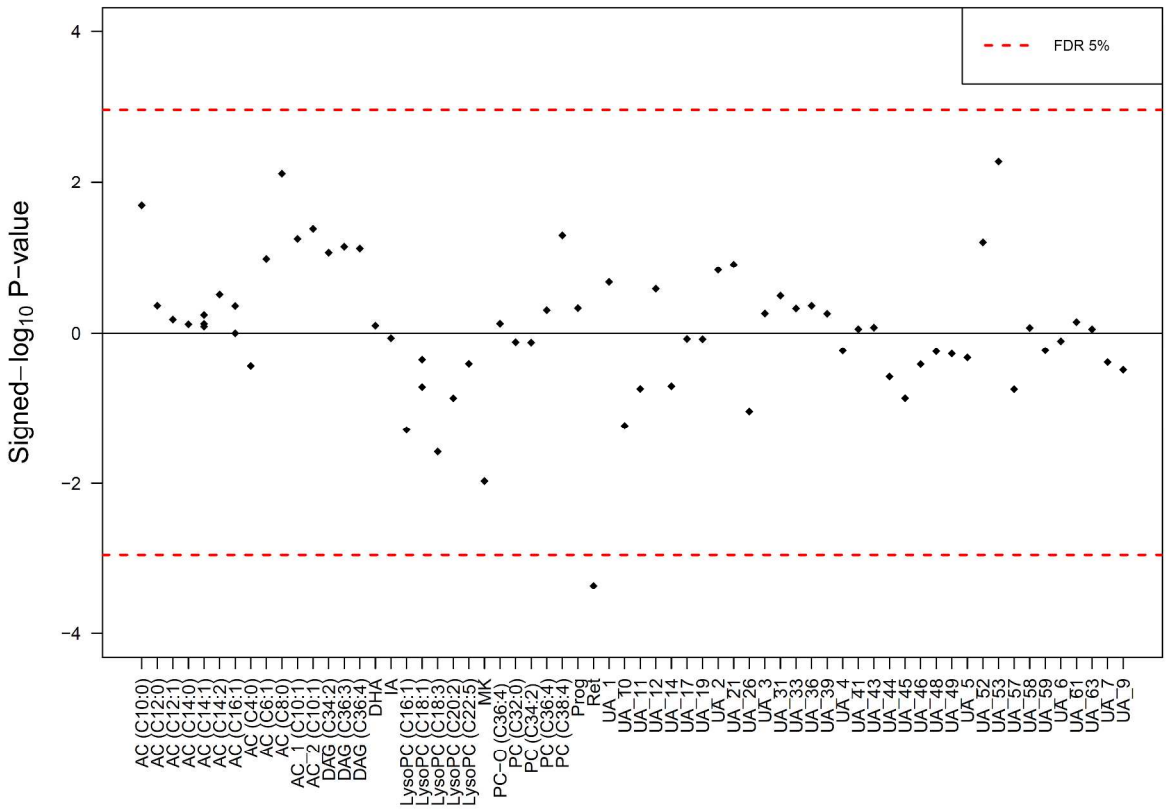


Figure 4: Manhattan Plot showing significance of association (log_{10} p value) with smoking (continuing into the 2nd trimester) for each birthweight-associated metabolite. Sign of p value indicates direction of association. Abbreviations: AC: acylcarnitine, UA: Unassigned, PC: Phosphatidylcholine, LysoPC: lysophosphatidylcholine, Prog: Progesterone, Ret: Retinol, DAG: Diacylglycerol, IA: Indolelactic acid, MK: methoxykynurenate, DHA: Docosahexaenoic acid. Numbers following unassigned compounds refer to compound numbers (see table s1).

Pathway analysis

Mummichog software assigned tentative annotations to 1629 of the 4714 features analysed and assigned tentative annotations to 30 of the 138 features associated with birthweight (additional dataset 2). Figure 5 shows significantly enriched pathways ($p < 0.05$) identified using *mummichog*. We observed enrichment of eight different pathways, two of these involve the synthesis and metabolism of signalling molecules, prostaglandins and steroid hormones, known to be involved in embryogenesis and child birth.

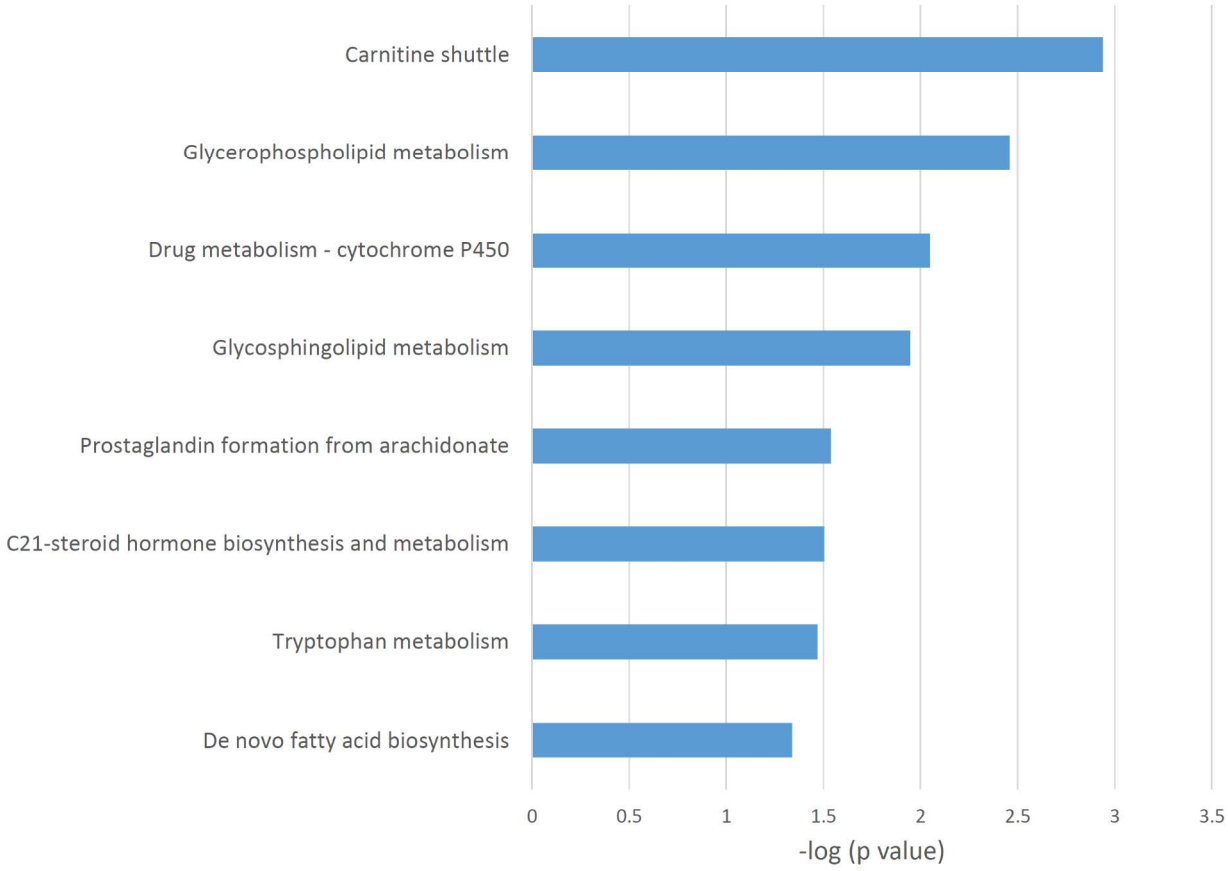


Figure 5: Metabolic Pathways significantly associated ($p < 0.05$) with birthweight in enrichment analysis.

Discussion

This is the largest study to date investigating untargeted metabolic profiles of cord blood associated with birth related outcomes, including weight at birth. Variance in birthweight, a trait that reflects the in uterine conditions throughout pregnancy and is an important predictor of health later in life, is determined by a complex combination of factors. Twin studies demonstrated the majority of these factors are of environmental origin[32]. Here, we have shown a metabolic signature associated with birthweight, after controlling for hereditary factors such as parental size, among healthy deliveries from the general European population. We observed changes in levels of vitamin A, progesterone, and molecules involved in pathways related to tryptophan metabolism, carnitine shuttle, fatty acid and glycerophospholipid metabolism.

Levels of vitamin A were higher with higher birthweight, confirming findings from previous studies[33-36]. Vitamin A likely promotes fetal growth through its role in cell proliferation and embryogenesis and interaction with nuclear receptors to alter gene expression[37]. While our results are consistent with changes in cord blood levels of vitamin A associated with smoking previously reported[38], this is the first study to the best of our knowledge that has demonstrated a link between vitamin A and both smoking and birth weight in the same population. Cigarette smoke has been shown to induce vitamin A depletion in animal models, and has been proposed to result from the induction of the CYP1A1 and CYP1A2 enzymes and the subsequent increase in catabolism of retinoic acid[39]. Vitamin A depletion therefore, may present an important aetiological pathway linking smoking with lower birthweight, which should be further investigated.

Progesterone levels were down-regulated with increasing birthweight. This corroborates findings from a recent targeted study of hormone levels in cord blood among Chinese and American births[40]. Progesterone is the major progestational hormone involved throughout all

stages of pregnancy. It is required for implantation, maintaining intrauterine conditions and initiating the signalling cascade to induce labour. Evidence for the relationship of progesterone measured earlier in pregnancy and birthweight is inconsistent[41-43]. However, levels of progesterone in cord blood appear unrelated to levels measured in maternal plasma earlier in pregnancy[44]. This reflects the transition in endocrine metabolism during fetal development, with the fetal endocrine system being well developed by late pregnancy. Pregnenolone, the precursor in the production of progesterone by the placenta, is transported to the fetus for metabolism of large quantities of C-19 steroid hormones by the fetal adrenal glands, which in turn signal back to regulate hormonal production by the maternal and placental systems[45]. The size of the fetus may influence production of molecules such as progesterone around the time of labour onset, for instance, to increase muscle contractibility during delivery. Furthermore, increased activity of the fetal endocrine system in larger babies may reduce availability of pregnenolone for progesterone production.

We observed a negative association with birthweight and methoxykynurenate and indolelactic acid. These metabolites, never previously associated with birthweight, are both final products of tryptophan metabolism, albeit through different routes, the kynurenine and indole pathways. Tryptophan itself was identified among the annotated features, and although there was no statistically significant association between the amino acid and birthweight ($p = 0.07$), our results indicate altered utilisation of tryptophan. Progesterone has been demonstrated to regulate tryptophan metabolism through inhibition of Trp 2,3-dioxygenase[46]. Tryptophan is essential throughout pregnancy firstly to meet the demand for protein synthesis during fetal development; to meet serotonin and kynurenine requirements; and to ultimately provide quinolinate for the production of NAD^+, which plays a key role in mitochondrial function[47]. All these requirements

will increase with greater fetal growth, thereby leaving less free tryptophan for conversion into methoxykynurenate and indolelactic acid. Similar results were reported by Favretto et al[48] who observed higher levels of tryptophan and lower levels of kynurenine in cord blood of intrauterine growth restricted babies. Animal experiments have showed cord blood levels of tryptophan to be related to maternal plasma levels[49] suggesting that cord blood levels may reflect tryptophan utilisation earlier in pregnancy. Therefore, monitoring of tryptophan metabolism throughout pregnancy may have utility in tracking the health of the developing fetus. Similarly, lower levels of multiple acylcarnitine species (C4, C6, C8, C10, C12, C14, and C16) were associated with increased birthweight, likely reflecting differences in energy utilization during development. Carnitine is an essential factor in fatty acid metabolism and its most important known metabolic function is to transport fatty acids into the mitochondria of cells for β-oxidation[50]. The placenta has a high activity of fatty acid oxidation enzymes[51] and where defects in long-chain fatty acid oxidation are noted, there is a higher frequency of small for gestation age babies[52]. Walsh et al[53] reported higher levels of acylcarnitines, including dodecanoylcarnitine identified in this study, in cord blood of infants asphyxiated during pregnancy, emphasising their importance to healthy fetal development. Clemente et al.[54] recently demonstrated in the INMA and ENVIR*ON*AGE cohorts that placental mitochondrial DNA content is associated with birthweight and may mediate the effects of environmental toxicants on birthweight. This supports our findings of the importance of metabolic pathways related to mitochondrial function. Mitochondria are particularly susceptible to oxidative stress and therefore may play a key role linking the fetal environment to growth.

The ω-3 fatty acid DHA has long been thought to be beneficial to fetal growth and recent randomized control trials have found that taking DHA supplements by pregnant women was

associated with increases in gestational length and birthweight[55-56]. Here, we observed a seemingly paradoxical negative association between cord blood DHA levels and birthweight. However, it has been reported that DHA intake by the mother explains only a small proportion of variance in DHA levels in cord blood[57], with the rest presumably explained by endogenous processes. Reduced DHA levels in cord blood would result from increased utilization in central nervous tissues and from greater eicosanoid metabolism. We found DHA levels to be associated with a cluster of metabolites tentatively assigned as diacylglycerols, which are molecules required for the production of arachidonic acid, an ω-6 fatty acid also involved in eicosanoid metabolism. Together these results highlight the role of long-chain fatty acid and potentially eicosanoid metabolism in fetal growth. Eicosanoids, in particular prostaglandins, are produced throughout pregnancy and play a role in regulating the maternal cardiovascular system, and like progesterone, are involved in signalling the onset of labour. Glycerophospholipids also showed significant disruption, indicating a variety of processes occurring in association with these molecular species. The lysoPCs were positively associated with birthweight and were negatively correlated with DHA and the diacylglycerols, suggesting some metabolic dependency. Positive associations between birthweight and lysoPCs have also been recently reported in a targeted analysis [17]. On the other hand, the PCs were associated with lower birthweight. This could reflect some cell membrane damage leading to the subsequent release of phospholipids, which have a variety of different proposed biological properties[58]. There is also recent evidence of antiphospholipid antibodies (and complement activation) co-operating in triggering a local inflammatory process[59] which may be linked to suboptimal fetal development.

We have complemented classical laboratory-based metabolic feature annotation in this study with the use of the recently developed *mummichog* algorithm to extract additional biological

information at the pathway level. Since the method matches metabolic features to pathways based on annotation through exact mass only, and does not account for any other physio-chemical identifiers, results of these analyses should be interpreted with caution. However, since feature misidentification likely applies equally to both the numerator (significant features) and denominator (total features detected) in enrichment analyses, the impact of wrong annotation may be less dramatic at the pathway level. We found that the majority of pathways identified, including tryptophan metabolism, carnitine shuttle, glycerophospholipid metabolism and C-21 steroid hormone biosynthesis, were supported by the laboratory annotation of metabolites.

Recent evidence indicates an effect of birthweight on later metabolic profile [60], with similarities to profiles associated with weight status measured at later life stages, suggesting that the metabolic profile of an individual at the start of life may persist into later life, with implications for health throughout the life course. Therefore, one may speculate that metabolic pathways identified in this study may provide a link between the observed associations with between birthweight and disease in adult life. In particular, *in utero* levels of progesterone could plausibly mediate the observed association between high birthweight and later development of breast cancer[6], as evidence is emerging that progesterone in adult life has a protective effect on breast cancer [61]. Conversely, tryptophan levels may play a role in the association between lower birthweight and increased risk of cardio-metabolic disease in later life, since tryptophan have been found to be predictive of subsequent development of type II diabetes in adult cohorts[62]. Following up current birth cohorts to adult life would be of great value to investigate these questions.

The main limitation of this study was related to its use of cord blood. Since samples were by necessity collected at the time of delivery, the study was cross-sectional in nature and therefore it

was difficult to disentangle whether perturbed metabolites were a cause or a result of variance in birthweight. However, cord blood provides a window into the direct supply of nutrients and other essential molecules to the developing fetus, and it also provides a snapshot of metabolism at the start of life. Our study was limited in scope to investigate changes specifically associated with low birthweight (<2500g) babies since we sampled from across the general population of births. However, the generalisability of our results, also considering the large number of samples collected from birth cohorts from across Europe, is a strength. Although any single metabolomics method cannot cover all the molecules of the cord blood metabolome, the UHPLC-MS platform we used represents a highly sensitive analytical technique able to measure hundreds to thousands of metabolites. Future work may include the incorporation of complementary metabolomics methods, other 'omic approaches such as DNA methylation analysis and multiplex analysis of cytokines, which are of interest due to their role in mediating hormonal signalling, and exploring the role of the cord blood metabolome in postnatal growth and development.

Conclusions

We have described metabolic profiles associated with birthweight among normal deliveries, highlighting the role of multiple metabolites in various pathways including tryptophan metabolism, fatty acid and glycerophospholipid metabolism and hormonal signalling. These results will have implications for antenatal and perinatal care, improve understanding of the pathways through which fetal growth may be affected and may have implications for health in later life.

SUPPORTING INFORMATION:

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

Figure S1. Directed acyclical graph to visualise assumptions regarding covariate, metabolome and birthweight relationships.

Figure S2. Whole metabolome principal component analysis

Table S1. All metabolomic features significantly associated with birthweight

Table S2. Stratified analysis by cohort, sex of baby and excluding caesarean and preterm (<37 weeks gestational age) deliveries

Table S3. Metabolite associations with birthweight, with adjustment for additional covariates.

Additional Dataset 1: Chromatograms and spectra.xls

Additional Dataset 2: Mummichog annotation.xls

AUTHOR INFORMATION

Corresponding Author

* Corresponding author: MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG. o.robinson@imperial.ac.uk

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally

Funding Sources

This work has been supported by the Exposomics EC FP7 grant (Grant agreement no: 308610). Data collection in the INMA Sabadell cohort was funded by grants from Instituto de Salud Carlos III (Red INMA G03/176) and the Generalitat de Catalunya (CIRIT 1999SGR 00241). OR was supported by an MRC Early Career Fellowship. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya. Piccolipiù was partially funded by the Italian National Centre for Disease Prevention and Control and by the Italian Ministry of Health.

REFERENCES

1. Seeds, J. W.; Peng, T., Impaired growth and risk of fetal death: is the tenth percentile the appropriate standard? *American journal of obstetrics and gynecology* **1998,** *178* (4), 658-69.

2. Ong, K. K.; Ahmed, M. L.; Emmett, P. M.; Preece, M. A.; Dunger, D. B., Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ (Clinical research ed.)* **2000,** *320* (7240), 967-71.

3. Goldenberg, R. L.; DuBard, M. B.; Cliver, S. P.; Nelson, K. G.; Blankson, K.; Ramey, S. L.; Herman, A., Pregnancy outcome and intelligence at age five years. *American journal of obstetrics and gynecology* **1996,** *175* (6), 1511-5.

4. Barker, D. J.; Osmond, C.; Kajantie, E.; Eriksson, J. G., Growth and chronic disease: findings in the Helsinki Birth Cohort. *Annals of human biology* **2009,** *36* (5), 445-58.

5. Nilsen, T. I.; Romundstad, P. R.; Troisi, R.; Potischman, N.; Vatten, L. J., Birth size and colorectal cancer risk: a prospective population based study. *Gut* **2005,** *54* (12), 1728-32.

6. Ahlgren, M.; Melbye, M.; Wohlfahrt, J.; Sorensen, T. I., Growth patterns and the risk of breast cancer in women. *The New England journal of medicine* **2004,** *351* (16), 1619-26.

7. Iniguez, C.; Ballester, F.; Costa, O.; Murcia, M.; Souto, A.; Santa-Marina, L.; Aurrekoetxea, J. J.; Espada, M.; Vrijheid, M.; Alvarez-Avellon, S. M.; Alvarez-Pedrerol, M.; Rebagliato, M., Maternal smoking during pregnancy and fetal biometry: the INMA Mother and Child Cohort Study. *American journal of epidemiology* **2013,** *178* (7), 1067-75.

8. Pedersen, M.; Giorgis-Allemand, L.; Bernard, C.; Aguilera, I.; Andersen, A. M. N.; Ballester, F.; Beelen, R. M. J.; Chatzi, L.; Cirach, M.; Danileviciute, A.; Dedele, A.; Eijsden, M. V.; Estarlich, M.; Fernández-Somoano, A.; Fernández, M. F.; Forastiere, F.; Gehring, U.; Grazuleviciene, R.; Gruzieva, O.; Heude, B.; Hoek, G.; Hoogh, K. D.; van den Hooven, E. H.; Håberg, S. E.; Jaddoe, V. W. V.; Klümper, C.; Korek, M.; Krämer, U.; Lerchundi, A.; Lepeule, J.; Nafstad, P.; Nystad, W.; Patelarou, E.; Porta, D.; Postma, D.; Raaschou-Nielsen, O.; Rudnai, P.; Sunyer, J.; Stephanou, E.; Sørensen, M.; Thiering, E.; Tuffnell, D.; Varró, M. J.; Vrijkotte, T. G. M.; Wijga, A.; Wilhelm, M.; Wright, J.; Nieuwenhuijsen, M. J.; Pershagen, G.; Brunekreef, B.; Kogevinas, M.; Slama, R., Ambient air pollution and low birthweight: A European cohort study (ESCAPE). *The Lancet Respiratory Medicine* **2013,** *1*, 695-704.

9. Govarts, E.; Nieuwenhuijsen, M.; Schoeters, G.; Ballester, F.; Bloemen, K.; de Boer, M.; Chevrier, C.; Eggesbo, M.; Guxens, M.; Kramer, U.; Legler, J.; Martinez, D.; Palkovicova, L.; Patelarou, E.; Ranft, U.; Rautio, A.; Petersen, M. S.; Slama, R.; Stigum, H.; Toft, G.; Trnovec, T.; Vandentorren, S.; Weihe, P.; Kuperus, N. W.; Wilhelm, M.; Wittsiepe, J.; Bonde, J. P., Birth

weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European Birth Cohorts. *Environmental health perspectives* **2012,** *120* (2), 162-70.

10. Fanos, V.; Atzori, L.; Makarenko, K.; Melis, G. B.; Ferrazzi, E., Metabolomics application in maternal-fetal medicine. *BioMed research international* **2013,** *9 pages, vol 2013. doi:10.1155/2013/72051*411. Horgan, R. P.; Broadhurst, D. I.; Walsh, S. K.; Dunn, W. B.; Brown, M.; Roberts, C. T.; North, R. A.; McCowan, L. M.; Kell, D. B.; Baker, P. N.; Kenny, L. C., Metabolic Profiling Uncovers a Phenotypic Signature of Small for Gestational Age in Early Pregnancy. *Journal of Proteome Research* **2011,** *10* (8), 3660-3673.

12. Maitre, L.; Fthenou, E.; Athersuch, T.; Coen, M.; Toledano, M. B.; Holmes, E.; Kogevinas, M.; Chatzi, L.; Keun, H. C., Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother-child cohort study. *BMC medicine* **2014,** *12*, 110.

13. Maitre, L.; Villanueva, C. M.; Lewis, M. R.; Ibarluzea, J.; Santa-Marina, L.; Vrijheid, M.; Sunyer, J.; Coen, M.; Toledano, M. B., Maternal urinary metabolic signatures of fetal growth and associated clinical and environmental factors in the INMA study. *BMC Medicine* **2016,** *14*, 177.

14. Dessì, A.; Atzori, L.; Noto, A.; Adriaan Visser, G. H.; Gazzolo, D.; Zanardo, V.; Barberini, L.; Puddu, M.; Ottonello, G.; Atzei, A.; Magistris, A. D.; Lussu, M.; Murgia, F.; Fanos, V., Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome. *The Journal of Maternal-Fetal & Neonatal Medicine* **2011,** *24* (sup2), 35-39.

15. Ivorra, C.; García-Vicent, C.; Chaves, F. J.; Monleón, D.; Morales, J. M.; Lurbe, E., Metabolomic profiling in blood from umbilical cords of low birth weight newborns. *Journal of Translational Medicine* **2012,** *10*, 142.

16. Tea, I.; Le Gall, G. n. l.; K??ster, A.; Guignard, N.; Alexandre-Gouabau, M. C. c.; Darmaun, D.; Robins, R. J., 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. https://doi.org/10.1371/journal.pone.0029947*

17. Hellmuth, C.; Uhl, O.; Standl, M.; Demmelmair, H.; Heinrich, J.; Koletzko, B.; Thiering, E., Cord Blood Metabolome Is Highly Associated with Birth Weight, but Less Predictive for Later Weight Development. *Obesity Facts* **2017,** *10* (2), 85-100.

18. Vineis, P.; Chadeau-Hyam, M.; Gmuender, H.; Gulliver, J.; Herceg, Z.; Kleinjans, J.; Kogevinas, M.; Kyrtopoulos, S.; Nieuwenhuijsen, M.; Phillips, D. H.; Probst-Hensch, N.; Scalbert, A.; Vermeulen, R.; Wild, C. P., The exposome in practice: Design of the EXPOsOMICS project. *International journal of hygiene and environmental health* **2017**, *Volume 220,* Issue 2, Part A, 142-151

19. Janssen, B. G.; Madlhoum, N.; Gyselaers, W.; Bijnens, E.; Clemente, D. B.; Cox, B.; Hogervorst, J.; Luyten, L.; Martens, D. S.; Peusens, M.; Plusquin, M.; Provost, E. B.; Roels, H. A.; Saenen, N. D.; Tsamou, M.; Vriens, A.; Winckelmans, E.; Vrijens, K.; Nawrot, T. S., Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study. *International journal of epidemiology* **2017**. *Volume 46,* Issue 5, 1 1386–1387m, https://doi.org/10.1093/ije/dyw269

20. Farchi, S.; Forastiere, F.; Vecchi Brumatti, L.; Alviti, S.; Arnofi, A.; Bernardini, T.; Bin, M.; Brescianini, S.; Colelli, V.; Cotichini, R.; Culasso, M.; De Bartolo, P.; Felice, L.; Fiano, V.; Fioritto, A.; Frizzi, A.; Gagliardi, L.; Giorgi, G.; Grasso, C.; La Rosa, F.; Loganes, C.; Lorusso, P.; Martini, V.; Merletti, F.; Medda, E.; Montelatici, V.; Mugelli, I.; Narduzzi, S.; Nistico, L.; Penna, L.; Piscianz, E.; Piscicelli, C.; Poggesi, G.; Porta, D.; Ranieli, A.; Rapisardi, G.; Rasulo, A.; Richiardi, L.; Rusconi, F.; Serino, L.; Stazi, M. A.; Toccaceli, V.; Todros, T.; Tognin, V.; Trevisan, M.; Valencic, E.; Volpi, P.; Ziroli, V.; Ronfani, L.; Di Lallo, D., Piccolipiu, a multicenter birth cohort in Italy: protocol of the study. *BMC pediatrics* **2014,** *14*, 36.

21. Guxens, M.; Ballester, F.; Espada, M.; Fernandez, M. F.; Grimalt, J. O.; Ibarluzea, J.; Olea, N.; Rebagliato, M.; Tardon, A.; Torrent, M.; Vioque, J.; Vrijheid, M.; Sunyer, J., Cohort Profile: the INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project. *International journal of epidemiology* **2012,** *41* (4), 930-40.

22. Chatzi, L.; Plana, E.; Daraki, V.; Karakosta, P.; Alegkakis, D.; Tsatsanis, C.; Kafatos, A.; Koutis, A.; Kogevinas, M., Metabolic syndrome in early pregnancy and risk of preterm birth. *American journal of epidemiology* **2009,** *170* (7), 829-36.

23. Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G.; de Hoogh, K.; Dedele, A.; Declercq, C.; Dimakopoulou, K.; Eeftens, M.; Forastiere, F.; Galassi, C.; Grazuleviciene, R.; Hoffmann, B.; Heinrich, J.; Iakovides, M.; Kunzli, N.; Korek, M.; Lindley, S.; Molter, A.; Mosler, G.; Madsen, C.; Nieuwenhuijsen, M.; Phuleria, H.; Pedeli, X.; Raaschou-Nielsen, O.; Ranzi, A.; Stephanou, E.; Sugiri, D.; Stempfelet, M.; Tsai, M. Y.; Lanki, T.; Udvardy, O.; Varro, M. J.; Wolf, K.; Weinmayr, G.; Yli-Tuomi, T.; Hoek, G.; Brunekreef, B., Evaluation of land use

regression models for NO2 and particulate matter in 20 European study areas: the ESCAPE project. *Environmental science & technology* **2013,** *47* (9), 4357-64.

24. Lazar, C. *imputeLCMD: A collection of methods for left-censored missing data imputation. R package*, 2.0; **2015**. https://CRAN.R-project.org/package=imputeLCMD

25. Campanella, G. *omics: '--omics' Data Analysis Toolbox. R package version* 0.1-4.; **2016.** https://CRAN.R-project.org/package=omics

26. R Core Team *R: A language and environment for statistical computing.*, Vienna, Austria. , **2016**. https://www.R-project.org/

27. Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A., HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic acids research* **2013,** *41* (Database issue), D801-7.

28. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN: a metabolite mass spectral database. *Therapeutic drug monitoring* **2005,** *27* (6), 747-51.

29. Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G., MyCompoundID: Using an Evidence-Based Metabolome Library for Metabolite Identification. *Analytical Chemistry* **2013,** *85* (6), 3401-3408.

30. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi,

R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R., Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007,** *3*, 211-221.

31. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F. H.; Khan, N.; Soltow, Q. A.; Jones, D. P.; Pulendran, B., Predicting Network Activity from High Throughput Metabolomics. *PLoS Computational Biology* **2013,** *9(7): e1003123. https://doi.org/10.1371/journal.pcbi.1003123*

32. Dubois, L.; Ohm Kyvik, K.; Girard, M.; Tatone-Tokuda, F.; Pérusse, D.; Hjelmborg, J.; Skytthe, A.; Rasmussen, F.; Wright, M. J.; Lichtenstein, P.; Martin, N. G., Genetic and environmental contributions to weight, height, and bmi from birth to 19 years of age: An international study of over 12,000 twin pairs. *PLoS ONE* **2012,** *(2): e30153. https://doi.org/10.1371/journal.pone.0030153*

33. Masters, E. T.; Jedrychowski, W.; Schleicher, R. L.; Tsai, W. Y.; Tu, Y. H.; Camann, D.; Tang, D.; Perera, F. P., Relation between prenatal lipid-soluble micronutrient status, environmental pollutant exposure, and birth outcomes. *The American journal of clinical nutrition* **2007,** *86* (4), 1139-45.

34. Rondo, P. H.; Abbott, R.; Tomkins, A. M., Vitamin A and neonatal anthropometry. *Journal of tropical pediatrics* **2001,** *47* (5), 307-10.

35. Gazala, E.; Sarov, B.; Hershkovitz, E.; Edvardson, S.; Sklan, D.; Katz, M.; Friger, M.; Gorodischer, R., Retinol concentration in maternal and cord serum: its relation to birth weight in healthy mother-infant pairs. *Early human development* **2003,** *71* (1), 19-28.

36. Ghebremeskel, K.; Burns, L.; Burden, T. J.; Harbige, L.; Costeloe, K.; Powell, J. J.; Crawford, M., Vitamin A and related essential nutrients in cord blood: relationships with anthropometric measurements at birth. *Early human development* **1994,** *39* (3), 177-88.

37. Clagett-Dame, M.; DeLuca, H. F., The role of vitamin A in mammalian reproduction and embryonic development. *Annual review of nutrition* **2002,** *22*, 347-81.

38. Chelchowska, M.; Ambroszkiewicz, J.; Gajewska, J.; Laskowska-Klita, T.; Leibschang, J., The effect of tobacco smoking during pregnancy on plasma oxidant and antioxidant status in mother and newborn. *European journal of obstetrics, gynecology, and reproductive biology* **2011,** *155* (2), 132-6.

39. Li, T.; Molteni, A.; Latkovich, P.; Castellani, W.; Baybutt, R. C., Vitamin A Depletion Induced by Cigarette Smoke Is Associated with the Development of Emphysema in Rats. *The Journal of Nutrition* **2003,** *133* (8), 2629-2634.

40. Lagiou, P.; Samoli, E.; Hsieh, C. C.; Lagiou, A.; Xu, B.; Yu, G. P.; Onoyama, S.; Chie, L.; Adami, H. O.; Vatten, L. J.; Trichopoulos, D.; Williams, M. A., Maternal and cord blood hormones in relation to birth size. *European Journal of Epidemiology* **2014,** *29*, 343-351.

41. Hartwig, I. R.; Pincus, M. K.; Diemert, A.; Hecher, K.; Arck, P. C., Sex-specific effect of first-trimester maternal progesterone on birthweight. *Human reproduction (Oxford, England)* **2013,** *28* (1), 77-86.

42. He, S.; Allen, J. C., Jr.; Malhotra, R.; Ostbye, T.; Tan, T. C., Association of maternal serum progesterone in early pregnancy with low birth weight and other adverse pregnancy outcomes. *The journal of maternal-fetal & neonatal medicine : the official journal of the*

European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet **2016,** *29* (12), 1999-2004.

43. Mucci, L. A.; Lagiou, P.; Tamimi, R. M.; Hsieh, C. C.; Adami, H. O.; Trichopoulos, D., Pregnancy estriol, estradiol, progesterone and prolactin in relation to birth weight and other birth size variables (United States). *Cancer causes & control : CCC* **2003,** *14* (4), 311-8.

44. Lagiou, P.; Hsieh, C. C.; Samoli, E.; Lagiou, A.; Xu, B.; Yu, G. P.; Onoyama, S.; Chie, L.; Vatten, L. J.; Adami, H. O.; Trichopoulos, D.; Williams, M. A., Associations of placental weight with maternal and cord blood hormones. *Annals of epidemiology* **2013,** *23* (11), 669-73.

45. Tal, R.; Taylor, H.; Burney, R.; al., e., Endocrinology of Pregnancy. [Updated 2015 Dec 7]. . In *Endotext [Internet]*, De Groot, L.; Chrousos, G.; Dungan, K.; al, e., Eds. MDText.com, Inc.: South Dartmouth (MA), 2000.Available from: https://www.ncbi.nlm.nih.gov/books/NBK278962/

46. Badawy, A. A., Effects of pregnancy on tryptophan metabolism and disposition in the rat. *The Biochemical journal* **1988,** *255* (1), 369-72.

47. Badawy, A. a.-B., The tryptophan utilization concept in pregnancy. *Obstetrics & gynecology science* **2014,** *57*, 249-59.

48. Favretto, D.; Cosmi, E.; Ragazzi, E.; Visentin, S.; Tucci, M.; Fais, P.; Cecchetto, G.; Zanardo, V.; Viel, G.; Ferrara, S. D., Cord blood metabolomic profiling in intrauterine growth restriction. *Analytical and Bioanalytical Chemistry* **2012,** *402*, 1109-1121.

49. Tsuji, A.; Nakata, C.; Sano, M.; Fukuwatari, T.; Shibata, K., L-Tryptophan Metabolism in Pregnant Mice Fed a High L-Tryptophan Diet and the Effect on Maternal, Placental, and Fetal Growth. *International Journal of Tryptophan Research : IJTR* **2013,** *6*, 21-33.

50. Borum, P. R., Carnitine in neonatal nutrition. *Journal of child neurology* **1995,** *10 Suppl 2*, S25-31.

51. Oey, N. A.; den Boer, M. E.; Ruiter, J. P.; Wanders, R. J.; Duran, M.; Waterham, H. R.; Boer, K.; van der Post, J. A.; Wijburg, F. A., High activity of fatty acid oxidation enzymes in human placenta: implications for fetal-maternal disease. *Journal of inherited metabolic disease* **2003,** *26* (4), 385-92.

52. Tyni, T.; Ekholm, E.; Pihko, H., Pregnancy complications are frequent in long-chain 3-hydroxyacyl-coenzyme A dehydrogenase deficiency. *American journal of obstetrics and gynecology* **1998,** *178* (3), 603-8.

53. Walsh, B. H.; Broadhurst, D. I.; Mandal, R.; Wishart, D. S.; Boylan, G. B.; Kenny, L. C.; Murray, D. M., The metabolomic profile of umbilical cord blood in neonatal hypoxic ischaemic encephalopathy. *PLoS One* **2012,** *7* (12), e50520.

54. Clemente, D. B.; Casas, M.; Vilahur, N.; Begiristain, H.; Bustamante, M.; Carsin, A. E.; Fernandez, M. F.; Fierens, F.; Gyselaers, W.; Iniguez, C.; Janssen, B. G.; Lefebvre, W.; Llop, S.; Olea, N.; Pedersen, M.; Pieters, N.; Santa Marina, L.; Souto, A.; Tardon, A.; Vanpoucke, C.; Vrijheid, M.; Sunyer, J.; Nawrot, T. S., Prenatal Ambient Air Pollution, Placental Mitochondrial DNA Content, and Birth Weight in the INMA (Spain) and ENVIRONAGE (Belgium) Birth Cohorts. *Environmental health perspectives* **2016,** *124* (5), 659-65.

55. Carlson, S. E.; Colombo, J.; Gajewski, B. J.; Gustafson, K. M.; Mundy, D.; Yeast, J.; Georgieff, M. K.; Markley, L. A.; Kerling, E. H.; Shaddy, D. J., DHA supplementation and pregnancy outcomes. *The American journal of clinical nutrition* **2013,** *97* (4), 808-15.

56. Harris, M. A.; Reece, M. S.; McGregor, J. A.; Wilson, J. W.; Burke, S. M.; Wheeler, M.; Anderson, J. E.; Auld, G. W.; French, J. I.; Allen, K. G., The Effect of Omega-3 Docosahexaenoic Acid Supplementation on Gestational Length: Randomized Trial of Supplementation Compared to Nutrition Education for Increasing n-3 Intake from Foods. *BioMed research international* **2015,** *vol. 2015, Article ID 123078, 8 pages, doi:10.1155/2015/123078*

57. Muhlhausler, B. S.; Gibson, R. A.; Yelland, L. N.; Makrides, M., Heterogeneity in cord blood DHA concentration: towards an explanation. *Prostaglandins, leukotrienes, and essential fatty acids* **2014,** *91* (4), 135-40.

58. Fruhwirth, G. O.; Loidl, A.; Hermetter, A., Oxidized phospholipids: from molecular properties to disease. *Biochimica et biophysica acta* **2007,** *1772* (7), 718-36.

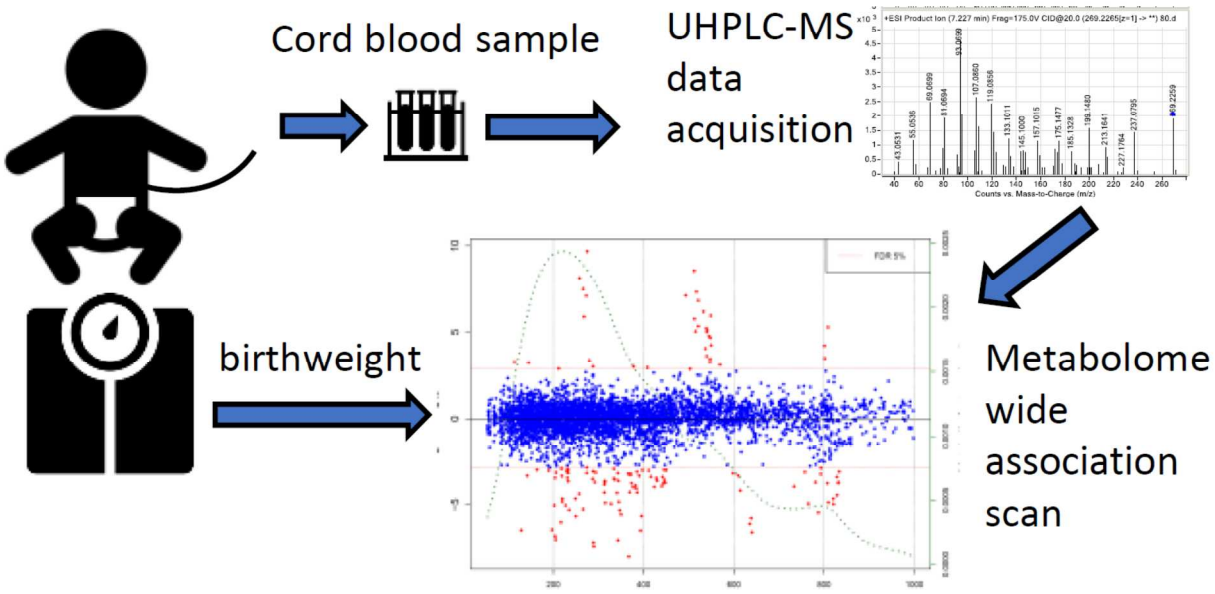
59. Tincani, A.; Cavazzana, I.; Ziglioli, T.; Lojacono, A.; De Angelis, V.; Meroni, P., Complement activation and pregnancy failure. *Clinical reviews in allergy & immunology* **2010,** *39* (3), 153-9.

60. Würtz, P.; Wang, Q.; Niironen, M.; Tynkkynen, T.; Tiainen, M.; Drenos, F.; Kangas, A. J.; Soininen, P.; Skilton, M. R.; Heikkilä, K.; Pouta, A.; Kähönen, M.; Lehtimäki, T.; Rose, R. J.; Kajantie, E.; Perola, M.; Kaprio, J.; Eriksson, J. G.; Raitakari, O. T.; Lawlor, D. A.; Davey Smith, G.; Järvelin, M.-R.; Ala-Korpela, M.; Auro, K., Metabolic signatures of birthweight in 18 288 adolescents and adults. *International journal of epidemiology* **2016,** *45*, 1539-1550.

61. Asi, N.; Mohammed, K.; Haydour, Q.; Gionfriddo, M. R.; Vargas, O. L.; Prokop, L. J.; Faubion, S. S.; Murad, M. H., Progesterone vs. synthetic progestins and the risk of breast cancer: a systematic review and meta-analysis. *Systematic reviews* **2016,** *5* (1), 121.

62. Chen, T.; Zheng, X.; Ma, X.; Bao, Y.; Ni, Y.; Hu, C.; Rajani, C.; Huang, F.; Zhao, A.; Jia, W.; Jia, W., Tryptophan Predicts the Risk for Future Type 2 Diabetes. *PLOS ONE* **2016,** *11* (9), e0162192.

For TOC only


Cord blood sample
UHPLC-MS
data
acquisition
birthweight
Metabolome
wide
association
scan