

# SCIENTIFIC REPORTS



OPEN

## Association between maternal exposure to phthalates and lower language ability in offspring derived from hair metabolome analysis

Beatrix Jones<sup>1</sup>, Ting-Li Han<sup>2,3</sup>, Thibaut Delplanck<sup>3</sup>, Elizabeth J. McKenzie<sup>1,2</sup>, Jamie V. de Seymour<sup>2</sup>, Mei Chien Chua<sup>4,5,6</sup>, Kok Hian Tan<sup>5</sup> & Philip N. Baker<sup>2,3,7</sup>

The fetus undergoes a crucial period of neurodevelopment in utero. The maternal hair metabolome provides an integrated record of the metabolic state of the mother prior to, and during pregnancy. We investigated whether variation in the maternal hair metabolome was associated with neurodevelopmental differences across infants. Maternal hair samples and infant neurocognitive assessments (using the Bayley III Scales of Infant Development at 24 months) were obtained for 373 infant-mother dyads between 26–28 weeks' gestation from the Growing Up in Singapore Towards Healthy Outcomes cohort. The hair metabolome was analysed using gas chromatography-mass spectrometry. Intensity measurements were obtained for 276 compounds. After controlling for maternal education, ethnicity, and infant sex, associations between metabolites and expressive language skills were detected, but not for receptive language, cognitive or motor skills. The results confirm previous research associating higher levels of phthalates with lower language ability. In addition, scores were positively associated with a cluster of compounds, including adipic acid and medium-chain fatty acids. The data support associations between the maternal hair metabolome and neurodevelopmental processes of the fetus. The association between phthalates and lower language ability highlights a modifiable risk factor that warrants further investigation.

Between 10–20% of individuals worldwide have a neurodevelopmental disability<sup>1</sup>. During pregnancy, the fetus undergoes a period of intense neurodevelopment, making this a key time frame for understanding variability in cognitive traits. Poor maternal nutrition is a known risk factor for neurodevelopmental delay in the infant<sup>2</sup>, as are environmental exposures<sup>3</sup>. The maternal metabolite profile, referred to as the metabolome, consists of the low molecular weight compounds (<1 KDa), and is dictated by the interplay between environment (including food intake and environmental exposures), the microbiome, and host cellular processes. In particular, studies have shown the maternal metabolome to be highly influenced by maternal BMI<sup>4</sup> and smoking<sup>5</sup>, as well as being responsive to the metabolic changes that occur in the development of pregnancy complications such as fetal growth restriction<sup>6</sup> and preeclampsia<sup>7</sup>. Targeted metabolomics studies have found links between the maternal plasma and urine metabolomes and infants' cognitive development<sup>8,9</sup>. Recently, maternal metabolomic profiling has expanded from analysis of typical biofluids such as blood and urine, and tissues such as placenta, to maternal hair. The maternal hair metabolome has been explored as a biological specimen to identify potential biomarkers and underlying metabolic mechanisms associated with the development of adverse pregnancy outcomes<sup>10,11</sup>. The hair metabolome has advantages over other more transient specimens such as urine and blood as it potentially

<sup>1</sup>Department of Statistics, University of Auckland, 38 Princes Street, Auckland, 1010, New Zealand. <sup>2</sup>Liggins Institute, University of Auckland, 85 Park Road, Grafton, Auckland, 1023, New Zealand. <sup>3</sup>First Affiliated Hospital of Chongqing Medical University, No. 1 Yixueyuan Road, Yuzhong District, Chongqing, 400016, China. <sup>4</sup>Department of Obstetrics & Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, NUHS Tower Block, Level 12, 1E Kent Ridge Road, Singapore, 119228, Singapore. <sup>5</sup>Department of Obstetrics and Gynaecology, KK Women's and Children's Hospital, 100 Bukit Timah Road, Singapore, 229899, Singapore. <sup>6</sup>DUKE-NUS Medical School, 8 College Road, Singapore, 169857, Singapore. <sup>7</sup>College of Life Sciences, University of Leicester, University Road, Leicester, LE1 7RH, United Kingdom. Correspondence and requests for materials should be addressed to J.V.d.S. (email: [j.deseymour@auckland.ac.nz](mailto:j.deseymour@auckland.ac.nz))

Participant Characteristics	Frequencies or Mean (sd)	P-values for associations with raw scores. For significant associations, Pearson's correlation (continuous variables) or the category with the highest mean score is given.				
		Cognitive	Receptive Language	Expressive Language	Fine Motor	Gross Motor
Ethnicity	Chinese: 58% Malay: 26% Indian: 16%	0.31	<b>0.02</b> Chinese	0.71	0.40	< <b>0.01</b> Malay
Highest Level of Maternal education	University: 39% GCE A level: 34% Secondary: 22% Primary: 5%	< <b>0.01</b> University	< <b>0.01</b> University	< <b>0.01</b> University	< <b>0.01</b> University	0.13
Sex of Child	Male: 58%	0.20	<b>0.01</b> Female	<b>0.02</b> Female	0.15	0.23
Child's age in days	733 (16)	<b>0.01</b> r = <b>0.14</b>	0.07	0.11	<b>0.00</b> r = <b>0.20</b>	0.99
Gestational Diabetes Mellitus (GDM)	Yes: 20%	<b>0.02</b> GDM	0.13	<b>0.02</b> GDM	<b>0.04</b> GDM	0.26
Smoking	Smoker: 40%	<b>0.03</b> Nonsmoker	< <b>0.01</b> Nonsmoker	0.09	0.24	0.83

**Table 1.** Participant characteristics significantly associated with at least one raw BSID-III score. Association p-values are based on Welch's t-test for dichotomous variables, ANOVA for categorical or ordinal variables with more than two categories, and tests of correlation for continuous variables. Values less than 0.05 are shown in bold. For significant associations, Pearson's correlation (continuous variables) or the category with the highest mean score is given.

offers a long term average of metabolite levels, and hair can be obtained non-invasively and stored easily. Pilot studies have suggested hair has a characteristic metabolite profile for type II diabetes<sup>12</sup> and animal studies have shown nutrition affects the amino acid composition of animal hair<sup>13</sup>. However, no studies have been reported to investigate associations between the maternal hair metabolome and neurocognitive development of the offspring. Given the associations reported between the maternal urine and plasma metabolome and infants' cognitive development, and the utility of the hair metabolome for identifying metabolomic markers of pregnancy disorders, the aim of our study was to investigate whether there are associations between the maternal hair metabolome and variation in the infant's neurocognitive development.

In this paper we focus on a subset of individuals from the GUSTO (Growing Up in Singapore Toward healthy Outcomes) cohort<sup>14</sup>. We have used gas-chromatography–mass spectrometry (GC-MS) to measure the metabolome of maternal hair collected at 26–28 weeks of pregnancy. We then assess whether there are associations between this metabolome and variation in infant neurocognitive development at age two years, evaluated using the five subscores of the Bayley III Scales of Infant Development (BSID-III)<sup>11,15</sup>: cognition, receptive language, expressive language, fine motor, and gross motor.

## Results

**Socio-demographic characteristics.** The socio-demographic characteristics of the 373 individuals included in this study that were significantly associated with one of the raw assessment scores are shown in Table 1. Ethnicity, sex of the child, age of the child at the time of cognitive assessment (measured in days), maternal education, gestational diabetes mellitus and smoking status were significantly associated with at least one cognitive outcome. Participant characteristics that were considered but were not found to be associated with any BSID-III raw scores were: maternal age, maternal BMI, gestational age at delivery (considered as a continuous variable, or as a dichotomous variable with categories before and after 37 weeks), and adjusted birth weight. Full results of all tests of association can be found in Supplementary Table 1. Many of the characteristics significantly associated with the cognitive scores were also associated with each other, so a more parsimonious model was constructed and used to control for socio-demographic characteristics in subsequent analyses. This model included ethnicity, sex of the child, age of the child at time of assessment, and maternal education. We note that explicitly controlling for the age of the child in days allows us to use the finer resolution raw score as a response.

**Metabolome.** Our untargeted analysis provided intensity measurements for 276 metabolites; 182 of these have putative (library match >75%) or tentative (library match between 60–75%) identifications.

The single metabolite analyses showed associations between some metabolites and the expressive language scale, controlling for the socio-demographic variables. The expressive language responses showed an excess of metabolites with small p-values (5% of p-values are expected to be less than 0.05 when all null hypotheses are true; the observed value was 17%), and moderate q-values (see Table 2). Cook's distance values were all less than 0.5, indicating that these relationships were not driven by outliers. Twenty-nine metabolites achieved our q-value criteria ( $q < 0.15$ ) for inclusion in multiple regression analyses. Two of these were then excluded based on the missing value criteria, leaving a group of 27 for further consideration. Table 3 shows the identified metabolites from this group, annotated with the direction of associations with expressive language score. Retention times for unidentified metabolites meeting the criteria are given in Supplementary Table 2.

Scale:	Cognitive	Receptive language	Expressive language	Fine Motor	Gross Motor
% pval < 0.05	8%	7%	17%	3%	9%
Min p-value	0.008	0.010	0.003	0.003	0.004
Min q-value	0.28	0.48	0.15	0.59	0.37

**Table 2.** Results of single metabolite analyses.

Positively Associated with Expressive Language Score		Negatively Associated with Expressive Language Score	
Metabolite	Univariate p-value	Metabolite	Univariate p-value
*Alanine Derivative	0.003	p-tert-Butylbenzoic acid (PTBBA)	0.009
*2-Bornanamine, N-methyl, peak 1	0.012	Dipicolinic acid	0.016
<b>Adipic Acid</b>	0.014	Homoalanine	0.019
Stearic acid	0.016	*Orthoacetic acid	0.022
<i>Nonanoic acid</i>	0.017	<b>Phthalic Acid</b>	0.022
Salicylic acid	0.022	Methionine	0.024
*2-Bornanamine, N-methyl, peak 2	0.023	Alpha-ketobutyric acid	0.026
*Nudifloric Acid	0.023		
<i>Decanoic Acid</i>	0.026		

**Table 3.** Identified metabolites achieving  $q < 0.15$  for the relationship with expressive language, and meeting the missing value criteria. Asterisks denote tentative identifications (library match between 60% and 75%). P-values are from the single metabolite analyses controlling for demographic factors. Bolded compounds were selected in the multivariate model; other compounds correlated to these selected compounds with  $|r| > 0.5$  are italicized.

**Multivariate predictor.** The selected multivariate multiple regression model for expressive language had 6 predictors (sex of the child, maternal education, ethnicity, and three metabolites) and an  $R^2$  of 0.15. Coefficient estimates can be found in Supplementary Table 3. The metabolites selected (Adipic acid, Phthalic acid, and an unidentified metabolite with retention time 12.59) are shown in bold in Table 3 and Supplementary Table 3. The model selection process is expected to retain a low number of representatives from groups of interrelated metabolites; those with moderately high correlations ( $r \geq 0.5$ ) to the selected compounds are particularly important for interpretation and are italicized in the tables.

The receiver operating curve for predicting scaled expressive language score  $\leq 5$  had an area under the curve of 0.71. This represents a modest improvement over the prediction using only demographic variables (AUC 0.67).

## Discussion

With intensities for 276 compounds, we have measured a more extensive hair metabolite profile than previous studies<sup>10–12</sup>. The hair profile contained compounds that we expected based on what is known about hair composition<sup>16</sup> and what types of compounds are detectable, given the choice of derivatisation method and analytical instrumentation. These comprised amino acids, benzenoids, carboxylic acids, fatty acids, hydroxy acids, and keto acids.

The raw expressive language scores from BSID-III have more and stronger associations with maternal hair metabolites than would be expected by chance, after adjusting for ethnicity, maternal education, exact infant age and sex. We note that we are examining ordinary variation in these scores, rather than disordered neurodevelopment; only the expected 2.5% (nine individuals) have their scaled expressive language scores  $< 4$  (i.e. more than two standard deviations below the mean using the BSID-III reference). The relationships observed provide intriguing links between the maternal hair metabolome and infants' neurocognitive development, and merit further investigation. While a large number of associations (29) met the q-value criteria, we have restricted our focus to metabolites that either appeared in the multivariate predictor as representatives of the relevant processes. To aid in interpretation, we also consider compounds correlated with the metabolites in the multivariate predictor with  $|r|$  at least 0.50.

In the multivariate model, adipic acid represents a group of metabolites that are positively associated with expressive language score, meaning that higher levels of adipic acid and its associated group were associated with higher expressive language scores on the BSD-III. Excretion of adipic acid is associated with ketosis<sup>17,18</sup> which can occur either when few carbohydrates are consumed, or when carbohydrate metabolism is impaired. Adipic acid is also included in controlled-release pharmaceuticals, antacids, and processed foods; it may be excreted when these are consumed<sup>19</sup>. In our sample, adipic acid is correlated at the 0.50 level with three compounds that also meet our q-value threshold: decanoic acid, nonanoic acid, and n-methyl-2-bornylamine. Decanoic acid is a major component of the medium chain triglyceride ketogenic anti-seizure diet; both decanoic and nonanoic acid have been shown to be effective for seizure control<sup>20</sup>. An investigation in *ex-vivo* rats has shown that the mechanism of seizure control for decanoic acid is inhibition of excitatory neurotransmission via AMPA glutamate receptor inhibition<sup>21</sup>.

AMPA receptors have been implicated in neuronal injury resulting from excitotoxicity<sup>22</sup>, particularly in the early postnatal period<sup>23</sup>. The association of AMPA inhibitors, nonanoic and decanoic acid, with higher expressive language scores poses an interesting question about whether these fatty acids have a role supporting optimal neurodevelopment.

Our analysis shows a negative relationship between language scores and phthalic acid levels, meaning that higher levels of phthalic acid were associated with lower language scores on the BSD-III – the opposite direction to the relationship observed between adipic acid and expressive language scores. Many studies have found associations between phthalic acid and a variety of cognitive disorders<sup>24</sup>. Phthalic acid is an endocrine disruptor, with sex-differentiated neurocognitive effects. Two independent studies found adverse effects on motor reflexes in male infants<sup>25,26</sup>. However, another study of infants aged 2–3 showed that girls' mental development index was negatively associated with phthalate exposure while boys' psychomotor development index was positively associated with phthalate exposure<sup>9</sup>. Other studies have shown that elevated phthalates in maternal urine are negatively associated with motor development for two and three year old girls<sup>5</sup>, and are positively associated with behavioral problems in eight year olds<sup>21</sup>. Phthalates are thought to affect steroidogenesis in the sex organs<sup>9,27,28</sup>, and while they definitely affect rodent brain development in utero<sup>29</sup>, little is known about their effect in utero in humans, although widespread disruptions in hippocampal functional and structural plasticity are postulated<sup>30,31</sup>. Interestingly, a Polish study found that prenatal phthalate exposure was inversely associated with child psychomotor development, but postnatal exposure did not<sup>32</sup>, and an American study found attention deficit in children whose mothers had high prenatal levels of urinary phthalates<sup>33</sup>.

Phthalates are common compounds added to plastics to make them flexible. Phthalic acid in biosamples originates from exposure to plastics, via food packaging, particularly of hot foods and beverages, dermal exposure to cosmetics, and inhalation<sup>34</sup>. Acute exposure can occur from prolonged contact with flexible medical tubing and catheters due to hospitalization; these exposures have been associated with long-term attention deficit<sup>35</sup>. Elevated phthalate levels in dust from a child's home have also previously been associated with lower verbal ability and developmental delay<sup>36</sup>. Phthalic acid was not correlated at the 0.50 level with any other known metabolites that met the q-value threshold.

The final compound in the multivariate model was an unknown with retention time 12.59 minutes. It is chemically similar to N-acetyl-1-methoxyethanimine (library match 56%). It did not have correlation 0.50 or larger with any known metabolites that met the q-value threshold.

While the non-invasive nature of hair sampling motivated the use of hair for this project, a shortcoming of our analysis was the use of full-length strands of hair. This means the metabolome measured was actually an average over the initial 26–28 weeks of pregnancy, as well as a variable period pre-pregnancy, depending on hair length. Future studies of the hair metabolome should consider performing segmental analysis, to allow the study of different time periods (e.g. pre-conception, first trimester, second trimester).

Another limitation was the cognitive assessment method. While BSID-III is widely used, it was developed in the United States and other researchers have found it necessary to establish new norms for different populations<sup>37</sup>. To check sensitivity to English language exposure, we considered a subset of our data ( $n = 257$ ) where the children were given a score rating their exposure to English. This score was not associated with the expressive language score ( $p = 0.31$ ), leading us to believe that the BSID-III was successfully administered in the Singaporean context. In addition, because we modelled the continuous scores rather than defining “low” or “high” categories, and adjusted for ethnicity in our models, we feel we have minimized the impact of cultural effects on our results.

It should also be taken into account that these findings are based on a cognitive assessment at a single time point. A conclusive finding of association between the maternal hair metabolome and the infants' neurocognitive development would require tracking infants' cognitive assessment over time. The findings from this study suggest such longitudinal tracking would be worthwhile.

In summary, the expressive language scores in the Bayley's assessment of infant development for infants at 24 months of age are associated with levels of several metabolites in the maternal hair metabolome. These associations were not seen for the receptive language, cognitive or motor scores. As in previous studies, higher maternal phthalate exposure was associated with poorer outcomes. This finding highlights an early modifiable risk factor for compromised infant neurodevelopment that warrants further investigation.

## Methods

**Ethical Statement.** The GUSTO study was approved by the National Health Care Group Domain Specific Review Board (reference D/09/021) and the SingHealth Centralized Institutional Review Board (reference 2009/280/D). Research was conducted according to the Declaration of Helsinki and all participants gave their written informed consent at recruitment.

**Study Participants.** The study was conducted on a subset of women from the GUSTO cohort - a large mother-child cohort that was first established in Singapore in June 2009. Women were recruited in their first trimester of pregnancy from KK Women's and Children's Hospital (KKH) and National University Hospital (NUH). Recruitment was completed in September 2010 ( $n = 1,247$ ). The cohort comprised three main ethnicities in the Singapore population (Chinese, Indian, and Malay). Of the 1,247 participants recruited in the GUSTO cohort, 373 mother-infant dyads had a neurodevelopmental assessment of the infant at two years of age, and a high quality maternal hair metabolome measurement.

**Cognitive assessment.** The Bayley III Scales of Infant Development (BSID-III)<sup>9</sup> were administered by trained assessors, fluent in the primary language of the child. We considered separately the five component scores: Cognitive, Receptive Language, Expressive Language, Fine Motor Skills and Gross Motor Skills. 365 children had all five scores, and 373 had at least one. Raw scores were used, with explicit modelling of the effect of the child's age in days.

Carrier Gas	Instrument grade helium (99.999%)
Sample injection	Automated injection 1 $\mu$ L
Injector liner	Deactivated glass split/splitless 4 mm ID straight single taper inlet liner packed with deactivated glass wool
Injector temperature	290 °C
Injector flow	Splitless, purge flow 25 mL/min, 1 min after injection
Column flow	1 mL/min, constant flow, column head pressure 9 psi
Column Type	Fused silica, 30 m $\times$ 250 $\mu$ m id $\times$ 0.15 $\mu$ m with 5 m guard column. Stationary phase of 86% dimethylpolysiloxane and 14% cyanopropylphenyl
Thermal program, transfer line and source temperatures, solvent delay	45 °C for 2 min, increased by 9 °C/min <sup>-1</sup> to 180 °C, held 5 min, then increased by 40 °C/min <sup>-1</sup> to 220 °C, held for 5 min, then increased by 40 °C/min <sup>-1</sup> to 240 °C, held for 11.5 min, increased by 40 °C/min <sup>-1</sup> to 280 °C and held for 2 min. The interface temperature was 250 °C and the quadrupole temperature was 230 °C.
Ionization mode	Positive, 70 eV
Acquisition mode	Operated in scan mode; started after 5.5 min with mass range between 38–550 amu with scan time of 0.1 s.
Detection Threshold	50 ion counts

**Table 4.** GCMS instrument parameters.

**Sample Preparation.** Hair samples were collected at 26–28 weeks of gestation. Hair was cut from the occipital area, 0.5 cm from the scalp with scissors. Samples were stored at room temperature until analysis. The hair samples were processed as part of a larger group of 971 samples. The order of hair samples was randomised, with the design across experimental days balanced with respect to gestational diabetes, small for gestational age, and preterm birth. Cognitive score was not considered when randomising the samples.

Hair samples were washed in Milli-Q water and methanol (Merck, analytical grade) twice. Hair strands were cut into short sections and weighed, and 0.5 mg–5.5 mg of hair (~2–3 strands) was spiked with 20  $\mu$ L of an internal standard mixture containing 10 mM of alanine-2,2,3,3-d<sub>4</sub>; 2 mM of hydroxyphenylalanine-3,3-d<sub>2</sub>; 10 mM of citric-2,2,4,4-d<sub>4</sub> acid; 10 mM of phenyl-d<sub>5</sub>-alanine (Sigma-Aldrich). Samples below 0.5 mg in weight were not included for analysis. Hair samples were hydrolysed by incubation in 1 mL of 1 M potassium hydroxide at 54 °C for 18 h. After hydrolysis, the hair extracts were neutralised to pH 7 by dropwise addition of sulphuric acid. To remove salts and other residues, 1 mL of methanol (100%v/v) was added to the neutralised samples and centrifuged at 2000 g for 5 min to separate the supernatant from insoluble hair debris. The supernatants containing hair extracts were evaporated to dryness using a rotary evaporator (Savant SPS121P Speedvac, Thermo) at 37 °C for 16 h. Samples were derivatised using methyl chloroformate (MCF)<sup>38</sup>. Briefly, samples were re-suspended in 200  $\mu$ L 1 M sodium hydroxide and transferred to a silanised glass tube. Methanol (167  $\mu$ L) and pyridine (34  $\mu$ L) were added. The derivatisation was initiated by adding 20  $\mu$ L MCF followed by vortexing for 30 s, then a further 20  $\mu$ L MCF was added, followed by 30 s of vortexing. To extract the derivisate, 400  $\mu$ L chloroform was added, vortexed for 10 s and then 400  $\mu$ L of 50 mM sodium bicarbonate was added and vortexed for an additional 10 s. The aqueous layer was discarded and the extract was dehydrated with anhydrous sodium sulphate before transferring to vials for GC-MS analysis.

The instrument used was an Agilent 7890B gas chromatograph coupled to an 5977 A single quadrupole mass spectrometer; settings are given in Table 4. Mass spectrometer calibration was carried out prior to analysis; the septum and liner were changed every ~200 injections. For each analytical batch of 18 samples, one solvent blank was run to measure instrument carryover, one alkane mix to measure instrument response, one reference standard mix to measure derivatisation efficiency, and one negative control to measure lab contamination. Five quality control samples (an aliquot pooled from every sample) were run throughout each batch to measure reproducibility.

**GC-MS Data Extraction.** The reference metabolite methyl chloroformate derivative mass spectral library developed by S. Villas-Boas (185 compounds) and the National Institute of Standards and Technology (NIST) 2014 EI library (242,477 compounds) were both used for compound identification. Using the NIST subset library method developed by E J McKenzie for the Automated Mass-spectral Deconvolution and Identification System (AMDIS), a subset of the NIST library was constructed using results from Chemstation Probability Based Matching integrator for a representative selection of quality control and sample data files. The Chemstation integration parameters were set to be sensitive to very low abundance compounds, and search-match parameters were expansive, with 10–15 compounds per feature included, in order to produce a comprehensive subset library of ~6000 compounds. This approach was used to circumvent limitations imposed by AMDIS on mass spectral library size.

The raw files from the GC-MS were converted into CDF/AIA format and then deconvoluted using AMDIS. An R-script<sup>39</sup> was used to integrate peak intensities. The R-script returns a value for the retention time bin defined by AMDIS for a particular ion across all samples, and is thus more sensitive than AMDIS, and has fewer missing values, but slightly more false positives. Both retention time and mass spectrum were used to match compounds against the Villas-Boas reference metabolite library. Mass spectrum alone was used to match against the NIST14 subset library. Matches below 60% are labelled as “unknown”. Matches between 60% and 75% are labelled as “tentative”, those >75% are putative identifications.

The values used were the maximum height at the apex of the most intense ion for the compound peak. Data were checked against negative controls to identify and remove analytical contaminants. Peaks that were not integrated well using the automated method were integrated individually using the R package Metab<sup>40</sup>.

**Normalisation.** Normalisation was performed using the 971 hair samples from study participants and 285 quality control samples. For each metabolite, the internal standard with the highest correlation  $r$  to that metabolite, in quality control samples, was used for normalisation, if  $r$  exceeded 0.85. Otherwise, internal standard normalisation was omitted. Samples were then normalised by biomass. Batch variation was addressed by median centering of the samples, specifically by multiplying the sample measurements for a particular day and metabolite by  $M/m$ , where  $m$  was the daily median and  $M$  the overall median for that metabolite. For a small number of days and metabolites, this created missing values if the median was a non-detection. These days were excluded for the study of that metabolite, and these metabolites were excluded from multi-metabolite analyses. For each metabolite, non-detections were replaced with 0.9 times the smallest non-zero value for that metabolite. Finally, values were log transformed.

**Statistical Analysis.** Associations between the five raw BSID-III scores and the characteristics in Table 2 were examined with correlation for continuous characteristics, Welch's t-test for dichotomous variables, and ANOVA for ordinal and categorical characteristics with more than two levels. From the characteristics with a significant ( $p < 0.05$ ) association for any of the five scores, a model was constructed to control for demographic factors in the single-metabolite regression analyses. This model included exact infant age, to allow for use of the raw scores, and enough demographic factors that adding further demographic variables did not significantly improve the fit of the model.

For each neurodevelopmental scale, a regression model was fitted, with the score as the response, and the chosen participant characteristics, along with a single log metabolite intensity, as the predictors. The  $p$ -values for the metabolic predictors were tabulated in each case, and  $q$ -values (estimates of false discovery rate) were computed using the method of Storey and Tibshirani<sup>41,42</sup>, separately for each neurocognitive score type.

Where there was an excess of significantly associated metabolites, we assessed the ability of the metabolites with  $q < 0.15$  to predict the BSID-III score using multivariate regression. We restricted our candidate predictors to the metabolites with complete information meeting the  $q$ -value criteria, and to participant characteristics in the selected demographic model for the neurodevelopmental scales. Only observations that were complete in all of these variables (365) were used for model selection. Backwards stepwise selection with AIC was used to produce a model, followed by stepwise pruning of predictors with the largest multivariate regression  $p$ -values until all  $p$ -values were less than 0.05. Using this model, receiver operator curves were produced for prediction of scaled expressive language score  $\leq 5$ .

To understand the relationships between the metabolite candidate predictors for the multiple regression, a correlation matrix was computed. Where correlation with a metabolite included in the selected model exceeded 0.5, these metabolites were considered as a group for purposes of interpretation.

## References

- Rudolph, A. & Lister, G. E. Ch. 11, (McGraw-Hill Professional, 2011).
- Grantham-McGregor, S. & Baker-Henningham, H. Review of the Evidence Linking Protein and Energy to Mental Development. *Public Health Nutr* **8**, 1191–1201 (2005).
- Rauh, V. A. & Margolis, A. E. Research Review: Environmental exposures, neurodevelopment, and child mental health - new paradigms for the study of brain and behavioral effects. *Journal of Chile Psychology and Psychiatry* **57**, 775–793 (2016).
- Sandler, V. et al. Associations of maternal BMI and insulin resistance with the maternal metabolome and newborn outcomes. *Diabetologia* **60**, 518–530, <https://doi.org/10.1007/s00125-016-4182-2> (2017).
- Rolle-Kampczyk, U. E. et al. Metabolomics reveals effects of maternal smoking on endogenous metabolites from lipid metabolism in cord blood of newborns. *Metabolomics* **12**, 76, <https://doi.org/10.1007/s11306-016-0983-z> (2016).
- Horgan, R. P. et al. Metabolic Profiling Uncovers a Phenotypic Signature of Small for Gestational Age in Early Pregnancy. *Journal of Proteome Research* **10**, 3660–3673, <https://doi.org/10.1021/pr2002897> (2011).
- Kenny, L. C. et al. Robust early pregnancy prediction of later preeclampsia using metabolomic biomarkers. *Hypertension (Dallas, Tex.: 1979)* **56**, 741–749, <https://doi.org/10.1161/hypertensionaha.110.157297> (2010).
- Wu, B. T., Dyer, R. A., King, D. J., Richardson, K. J. & Innis, S. M. Early second trimester maternal plasma choline and betaine are related to measures of early cognitive development in term infants. *PLoS One* **7**, 3448 (2012).
- Télez-Rojo, M. M. et al. Prenatal urinary phthalate metabolites levels and neurodevelopment in children at two and three years of age. *Science of The Total Environment* **461–462**, 386–390, <https://doi.org/10.1016/j.scitotenv.2013.05.021> (2013).
- He, X. et al. Maternal hair metabolome analysis identifies a potential marker of lipid peroxidation in gestational diabetes mellitus. *Acta diabetologica* **53**, 119–122, <https://doi.org/10.1007/s00592-015-0737-9> (2016).
- Sulek, K. et al. Hair metabolomics: identification of fetal compromise provides proof of concept for biomarker discovery. *Theranostics* **4**, 953–959, <https://doi.org/10.7150/thno.9265> (2014).
- Rashaid, A. H., Harrington, P. D. B. & Jackson, G. P. Profiling Amino Acids of Jordanian Scalp Hair as a Tool for Diabetes Mellitus Diagnosis: A Pilot Study. *Analytical Chemistry* **87**, 7078–7084 (2015).
- Almeida, A. M. et al. Influence of feed restriction on the wool proteome: a combined iTRAQ and fiber structural study. *Journal of Proteomics* **103**, 170–177 (2014).
- Soh, S.-E. et al. Insights from the Growing Up in Singapore Toward Healthy Outcomes (GUSTO) Cohort Study. *Ann Nutr Metab* **64**, 218–225, <https://doi.org/10.1159/000365023> (2014).
- Bayley, N. *Bayley Scales of Infant and Toddler Development, Third Edition: Administration Manual*. Third edn, (PsychCorp, 2006).
- Yu, J., Yu, D., Checkla, D., Freedberg, I. & Bertolino, A. Human Hair Keratins. *Journal of Investigative Dermatology* **101**, 56S–59S (1993).
- Liebich, H. M., Pickert, A., Stierle, U. & Woll J. Gas chromatography-mass spectrometry of saturated and unsaturated dicarboxylic acids in urine. *Journal of Chromatography A* **199**, 181–189 (1980).
- Mortensen, P. B. The possible antiketogenic and gluconeogenic effect of the omega oxidation of fatty acids in rats. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **620**, 177–185 (1980).
- Glick, N. & Fischer, M. Adipic aciduria, a dietary artefact. *Journal of Inherited Metabolic Disease* **14**, 849–850 (1991).

20. Chang, P. *et al.* Seizure control by ketogenic diet associated medium chain fatty acids. *Neuropharmacology* **69**, 105–114 (2013).
21. Chang, P. *et al.* Seizure control by decanoic acid through direct AMPA receptor inhibition. *Brain Development* **139**, 431–443 (2016).
22. Burd, I., Welling, J., Kannan, G. & Johnston, M. V. Excitotoxicity as a Common Mechanism for Fetal Neuronal Injury with Hypoxia and Intrauterine Inflammation. *Advances in pharmacology (San Diego, Calif.)* **76**, 85–101, <https://doi.org/10.1016/bs.apha.2016.02.003> (2016).
23. McDonald, J. W., Trescher, W. H. & Johnston, M. V. Susceptibility of brain to AMPA induced excitotoxicity transiently peaks during early postnatal development. *Brain research* **583**, 54–70 (1992).
24. Ejaredar, M., Nyanza, E. C., Ten Eycke, K. & Dewey, D. Phthalate exposure and childrens neurodevelopment: A systematic review. *Environmental Research* **142**, 51–60, <https://doi.org/10.1016/j.envres.2015.06.014> (2015).
25. Engel, S. M. *et al.* Prenatal phthalate exposure and performance on the neonatal behavioral assessment scale in a multiethnic birth cohort. *Neurotoxicology* **30**, 522–528, <https://doi.org/10.1016/j.neuro.2009.04.001> (2009).
26. Yolton, K. *et al.* Prenatal exposure to bisphenol A and phthalates and infant neurobehavior. *Neurotoxicology and Teratology* **33**, 558–566, <https://doi.org/10.1016/j.ntt.2011.08.003> (2011).
27. Svechnikov, K. *et al.* Phthalates Exert Multiple Effects on Leydig Cell Steroidogenesis. *Hormone Research in Paediatrics* **86**, 253–263 (2016).
28. Cowell, W. J. & Wright, R. J. Sex-Specific Effects of Combined Exposure to Chemical and Non-chemical Stressors on Neuroendocrine Development: a Review of Recent Findings and Putative Mechanisms. *Curr Environ Health Rep* **4**, 415–425, <https://doi.org/10.1007/s40572-017-0165-9> (2017).
29. Lin, H. *et al.* In Utero Exposure to Diethylhexyl Phthalate Affects Rat Brain Development: A Behavioral and Genomic Approach. *International Journal of Environmental Research and Public Health* **12**, 13696 (2015).
30. Ponsonby, A.-L. *et al.* Epigenetic regulation of neurodevelopmental genes in response to in utero exposure to phthalate plastic chemicals: How can we delineate causal effects? *Neurotoxicology* **55**, 92–101, <https://doi.org/10.1016/j.neuro.2016.05.011> (2016).
31. Holahan, M. R. & Smith, C. A. Phthalates and neurotoxic effects on hippocampal network plasticity. *Neurotoxicology* **48**, 21–34, <https://doi.org/10.1016/j.neuro.2015.02.008> (2015).
32. Polanska, K., Ligońka, D., Sobala, W. & Hanke, W. Phthalate exposure and child development: The Polish Mother and Child Cohort Study. *Early Human Development* **90**, 477–485, <https://doi.org/10.1016/j.earlhumdev.2014.06.006> (2014).
33. Factor-Litvak, P. *et al.* Persistent Associations between Maternal Prenatal Exposure to Phthalates on Child IQ at Age 7 Years. *PLOS ONE* **9**, e114003, <https://doi.org/10.1371/journal.pone.0114003> (2014).
34. Sidlovska, M. *et al.* Exposure of children to phthalates and the impact of consumer practices in Slovakia. *Reviews on environmental health* **32**, 211–214, <https://doi.org/10.1515/reveh-2016-0028> (2017).
35. Verstraete, S. *et al.* Circulating phthalates during critical illness in children are associated with long-term attention deficit: a study of a development and a validation cohort. *Intensive Care Medicine* **42**, 379–392, <https://doi.org/10.1007/s00134-015-4159-5> (2016).
36. Phillipat, C. *et al.* Phthalate concentrations in house dust in relation to autism spectrum disorder and developmental delay in the CHildhood Autism Risks from Genetics and the Environment (CHARGE) study. *Environmental Health* **14**, 56 (2015).
37. Cromwell, E. A. *et al.* Validity of US norms for the Bayley Scales of Infant Development-III in Malawian Children. *Eur J Paediatr Neurol* **18**, 223–230 (2014).
38. Smart, K. F., Aggio, R. B., Van Houtte, J. R. & Villas-Boas, S. G. Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography-mass spectrometry. *Nature Protocols* **5**, 1709–1729 (2010).
39. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna Austria, 2015).
40. Aggio, R., Villas-Boas, S. G. & Ruggiero, K. Metab: an R package for high throughput analysis of metabolomics data generated by GC-MS. *Bioinformatics* **27**, 2316–2318 (2011).
41. qvalue: Q-value estimation for false discovery rate control. R package version 2.0.0. v. 2.0.0 (2015).
42. Storey, J. & Tibshirani, R. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9440–9445 (2003).

## Acknowledgements

Funding was provided by the Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A\*STAR), Singapore; and the Ministry of Business, Innovation, and Employment (MBIE), New Zealand, through the International Relationships Fund. This research was also supported by the Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Programme and administered by the Singapore Ministry of Health's National Medical Research Council (NMRC), Singapore-NMRC/TCR/004-NUS/2008; NMRC/TCR/012-NUHS/2014; and the Bill and Melinda Gates Foundation (Grand Challenges Explorations, OPP1119254). Mass spectrometry analysis was carried out at the Centre for Genomics, Proteomics and Metabolomics at the University of Auckland. GC-MS technical assistance was provided by Erica Zarate and EJM. The authors thank Silas Villas Boas, for insightful discussions. We gratefully acknowledge Dr. Anne Rifkin-Graboi and her lab at A\*STAR for providing the neurocognitive assessment data and very helpful comments on an early draft of the manuscript, and Dr. Shaun Kok Yew Goh of National University of Singapore, Department of Biomedical Engineering, who provided the assessment of English language exposure.

## Author Contributions

P.N.B. conceived the study. T.D. performed the experiment under the guidance of T.-L.H., T.-L.H. developed the software used for data processing. T.D. and J.V.D.S. extracted the data under the guidance of E.J.M. and T.-L.H., B.J. carried out the statistical analysis and B.J., E.J.M. and J.V.D.S. interpreted the results. B.J. wrote the manuscript with help from E.J.M., J.V.D.S., and T.D. All authors revised the manuscript prior to submission.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-24936-5>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018