Early Life Environment

Prenatal mercury concentration is associated with changes in DNA methylation at \textit{TCEANC2} in newborns

Kelly M Bakulski,1† HwaJin Lee,2† Jason I Feinberg,1,2 Ellen M Wells,3 Shannon Brown,1 Julie B Herbstman,1,4 Frank R Witter,2 Rolf U Halden,2,5 Kathleen Caldwell,6 Mary Ellen Mortensen,6 Andrew E Jaffe,2,7 John Moye Jr,6 Laura E Caulfield,1 Yi Pan,6 Lynn R Goldman,1,9‡ Andrew P Feinberg1,2‡ and M Daniele Fallin1,2*‡

1Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA, 2Johns Hopkins University School of Medicine, Baltimore, Maryland, USA, 3Purdue University, School of Health Sciences, West Lafayette, Indiana, USA, 4Columbia University Mailman School of Public Health, New York City, New York, USA, 5Arizona State University, Fulton School of Engineering, Tempe, Arizona, USA, 6National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Georgia, USA, 7Lieber Institute for Brain Development, Baltimore, Maryland, USA, 8Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland, USA and 9George Washington University School of Public Health, Washington D.C., USA

*Corresponding author. 624 N. Broadway, Hampton House 850, Baltimore, Maryland 21205, US. E-mail: dfallin@jhu.edu.
†These authors are joint first authors.
‡These authors contributed equally as senior leaders of this project.

Accepted 26 February 2015

Abstract

Background: Human exposure to the widespread environmental contaminant mercury is a known risk factor for common diseases such as cancer, cardiovascular disease and neurological disorders through poorly characterized mechanisms. Evidence suggests mercury exposure may alter DNA methylation levels, but to date, the effects in early life on a genome-wide scale have not been investigated.

Methods: A study sample of 141 newborns was recruited in Baltimore, MD, USA and total mercury and methylmercury were measured in cord blood samples. We quantified genome-wide DNA methylation data using CHARM 2.0, an array-based method, and used region-finding analyses to identify concentration-associated differentially methylated regions (DMRs). To test for replication of these identified DMRs in the pilot, or Vanguard, phase of the National Children’s Study (NCS), we compared bisulfite-pyrosequenced DNA at candidate regions from 85 whole cord blood samples with matched first trimester maternal mercury concentration measures.

Results: Total mercury concentration was associated with methylation at DMRs inside \textit{ANGPT2} and near \textit{PRPF18} genes [false discovery rate (FDR) < 0.05], as well as DMRs near...
FOXD2 and within TCEANC2 (FDR < 0.1) genes. Methylmercury concentration was associated with an overlapping DMR within TCEANC2 (FDR < 0.05). In NCS replication analyses, methylation levels at three of four cytosine-guanine DNA dinucleotides (CpG sites) within the TCEANC2 DMR were associated with total mercury concentration (P < 0.05), and this association was diminished after adjusting for estimated cell proportions.

**Conclusions:** Evidence for an association between mercury and DNA methylation at the TCEANC2 region was found, which may represent a mercury-associated shift in cord blood cell composition or a change in methylation within blood cell types. Further confirmatory studies are needed.

**Key words:** DNA methylation, mercury exposure, epigenetics, cord blood, methylmercury

---

**Key Messages**

- *In utero* mercury exposure may be related to DNA methylation changes in offspring.
- In a genome-wide DNA methylation analysis of 141 cord blood samples, four differentially methylated regions were associated with total or methylmercury concentrations at false discovery rate q-values < 0.1.
- The region associated with the TCEANC2 gene was validated in the same samples using a second method of DNA methylation detection, and it was replicated in a second independent sample of infants.

---

**Introduction**

Environmental and occupational mercury exposure has long been recognized to have severe health consequences, including sensory disturbance, visual field constriction, ataxia, deafness and neurodevelopmental deficits. Human toxicity to mercury depends on the mercury species, including methylmercury, ethylmercury and inorganic mercury, all of which are encompassed in blood measures of total mercury. Typical exposure is via dietary methylmercury from fish and other seafood. Toxicity from dietary methylmercury has been observed historically, including the Minamata, Japan and Iraq grain poisoning epidemics, where very high methylmercury in utero exposures caused severe central nervous system damage and death. Much lower methylmercury in utero exposures are also associated with subtle, subclinical brain function deficits. The mechanisms underlying the early life health effects of mercury exposure are incompletely characterized and suggestions include localized neuronal death, microtubule problems, neuronal migration disruption, oxidative damage and altered brain calcium metabolism. Given the potential for epigenetic susceptibility to toxicant exposures, fetal epigenetic changes resulting from in utero mercury exposure could inform this mechanism of toxicity.

Epigenetic epidemiology research, particularly genome-wide association studies (EWAS), interrogates epigenetic changes associated with environmental exposures or disease phenotypes in a population context. Identified epigenetic markers may act directly in disease causality, or they may interact with genetic sequence variation or environmental factors to influence disease. Alternatively, they may be useful biomarkers of previous exposures when the toxicant and metabolites may no longer be observable. The prenatal period is a critical window in development and the potential relationship between mercury body burden and DNA methylation at birth may provide an opportunity for epidemiological research with implications concerning central nervous system development.

At present, no genome-scale study has tested the association between DNA methylation and mercury exposure during prenatal development. We performed an genome-wide association study of cord blood DNA methylation in 141 newborns to identify regions ofDNA methylation correlated with in utero mercury exposure. We then tested for replication of the identified regional DNA methylation-mercury association results in an independent sample, using a second method of DNA methylation detection. The findings from this genome-scale study enhance our understanding of the relationship between heavy metal exposures and DNA methylation early in fetal development, with potential implications for the usefulness of epigenetic changes as biomarkers of prenatal exposure and potential mediators of health effects.
Methods

Study populations

Discovery sample

Cord blood samples were obtained from newborns at Johns Hopkins Hospital, MD, USA, between November 2004 and March 2005, as part of the previously described cross-sectional Baltimore Tracking Health Related to Environmental Exposures (THREE) study. Demographic and limited medical information were obtained from maternal and neonatal medical records. Study clinicians reviewed a 10% random sample of medical record abstractions for accuracy, as previously described. The THREE study was reviewed and approved by the Johns Hopkins School of Medicine Institutional Review Board (IRB). A waiver of the need for informed consent was obtained since the study constituted minimal risk, was anonymous and made use of only biological samples that would otherwise be medical waste.

During the Baltimore THREE study time window, there were 609 live births. Twelve twin births (24 infants) and 256 births where cord blood was not available were excluded. An additional 41 births had insufficient cord blood samples for methylmercury analyses, leaving 300 infants eligible for the current study. Among the study samples, 187 had umbilical cord blood clots available for DNA extraction, 167 samples yielded sufficient DNA for the methylation assay (5 μg) and 141 samples passed our internal DNA methylation array quality control. Samples in the THREE epigenetics set (n = 141) were similar to the samples in the larger THREE exposure study (n = 300) with respect to maternal age, race, mercury exposure, infant sex, gestational age and birthweight (Supplementary Table 1, available as Supplementary data at IJE online).

Replication sample

Maternal matched first and third trimester peripheral blood samples (nwomen = 147), and a subset of corresponding cord blood samples (n = 90) were obtained from the Initial Vanguard Centers of the National Children’s Study, a pilot study involving mothers and infants recruited from seven different locations across the USA (Duplin County, NC; Queens County, NY; Orange County, CA; Waukesha County, WI; Salt Lake County, UT; Montgomery County, PA; and, a composite location of four adjacent counties in South Dakota and Minnesota). The NCS is an ongoing, large, long-term study of children’s health and development that has previously been described. The Vanguard phase of the NCS began in 2009, in order to inform the later Main Study, to evaluate the methodologies, feasibility, acceptability and cost of the study. Women were eligible if they were biologically able to become pregnant or were currently pregnant and were aged 18 through 49 years. Participants had up to two pregnancy study visits in the first and third trimesters and a visit with their infants at birth. Demographic information was collected by interview at study visits before or during pregnancy. Race/ethnicity categories of mothers were self-reported as non-Hispanic White, non-Hispanic Black, and all Hispanics including Mexican-Americans and other groups. Due to limited sample size, these were re-categorized as White/non-White for all methylation analyses. Written informed consent was obtained from all NCS participants or their legally authorized representatives and the protocol was approved by the corresponding study centre Institutional Review Board. The infants were born between October 2009 and June 2011. Only material from participants who explicitly consented to use of their biological samples to obtain genetic information was used in these analyses. Of the 90 cord blood samples, 85 had sufficient DNA for replication methylation analyses; however, fewer had available exposure data as described below.

Exposure assessments

In the THREE study samples, total mercury, methylmercury, lead, selenium, copper and omega-3 fatty acids were analysed as previously described. Umbilical cord whole blood and serum samples were collected in trace element-free tubes. Whole blood total mercury and lead were measured via inductively coupled plasma mass spectrometry (ICP-MS) and serum concentrations of selenium and copper were measured using inductively coupled plasma dynamic reaction cell mass spectrometry (ICP-DRC-MS) at the Centers for Disease Control and Prevention (CDC) National Center for Environmental Health (CDC/NCEH) laboratory. The CDC’s Environmental Health Laboratory was determined not to be participating in human subjects research. The cord serum n-3 polyunsaturated acids, eicosapentanoic acid (EPA) and docosahexaenoic acid (DHA) were measured using automated fast gas chromatography at the National Institute for Alcohol Abuse and Alcoholism and their levels were summed for analysis (EPA + DHA). The limits of detection (LOD) for exposure concentrations in the THREE study were 0.48 μg/l (methylmercury), 0.33 μg/l (total mercury), 0.25 μg/dl (lead), 5 μg/l (selenium) and 4 μg/dl (copper). Samples with exposure values below the LOD were treated as LOD/√2, except for methylmercury. As there were more than 10% values < LOD for methylmercury, these were replaced with concentrations reported from the laboratory instrument, in order to minimize bias due to left-censoring.
In the NCS Vanguard study, mercury measures in infant cord blood samples were not available. Instead, blood samples collected from pregnant mothers in their first trimester were used for exposure measurement. Blood was collected into CDC-prescreened EDTA tubes that were free of lead, mercury, selenium and copper. Whole blood samples were frozen and shipped on dry ice to the NCS Repository (Fisher Bioservices, Rockville, MD), then stored frozen at vapour phase liquid nitrogen temperatures (between −135 and −190°C) until shipment on dry ice to the CDC laboratory where they were also stored frozen (−70°C) until they were analysed. The NCEH measured whole blood total mercury, lead and selenium via ICP-MS in the NCS samples as well (LODs respectively: 0.28 total Hg μg/l, 0.25 Pb μg/dl, 5 Se μg/l). Methylmercury was measured via triple-spike isotope dilution with solid-phase microextraction gas chromatography (TSID-SPME-GC) (LOD 0.12 μg/l).19 Of the maternal samples, 77 had available methylmercury measures and corresponding infant DNA methylation data from targeted pyrosequencing assays (see below); 75 had available total mercury data and pyrosequencing DNA methylation data. For cell-type association estimates to explore the potential for confounding (see Statistical methods), 138 mothers had mercury measures and 450K array DNA methylation data (see below).

DNA methylation measures

Genome-scale DNA methylation discovery: CHARM 2.0 DNA methylation in the THREE study samples was measured via the second version of comprehensive high-throughput arrays for relative methylation (CHARM 2.0)(Roche NimbleGen, Madison, WI, USA).13 These measures were used to perform genome-scale discovery of candidate differentially methylated regions (DMRs) associated with total mercury and methylmercury concentrations. DNA was isolated from THREE study cord blood clot samples using DNeasy Blood & Tissue kits (Qiagen, Valencia, CA, USA) following the manufacturer’s instructions. Of 187 cord blood clot samples, 167 (89.3%) yielded sufficient DNA. After DNA isolation, 5 μg of purified genomic DNA was subjected to shearing, digestion, purification, amplification, labelling and hybridization as previously described,20 generating data on 2.1 million probes covering 5.2 million CpGs, including all CpG islands, CpG island shores, annotated and non-annotated promoters, and microRNA sites. At the sample level, 26 individuals were excluded because observed intensities of < 80% of probes exceeded background levels. At the probe level, sex chromosome probes and those with signals below background in > 25% of samples were removed. The final analysis included 141 samples with 1 569 888 autosomal probes covering 4 254 946 CpGs.

The array hybridization and subsequent processing were performed over 5 days; we addressed this potential batch effect by surrogate variable analysis (SVA) as described in the Statistical methods subsection. Total and methylmercury concentrations did not vary by array hybridization batch (P-value<sub>total</sub> = 0.58; P-value<sub>methyl</sub> = 0.25).

Pyrosequencing validation of CHARM in THREE samples DNA methylation, at the candidate DMRs (in or near TCEANC2, ANGPT2, PRPF18 and FOXD2 genes) identified during genome-wide discovery phase as surpassing the significance threshold, were validated via bisulfite pyrosequencing on THREE study samples for which additional bisulfite converted DNA was available (n = 136). MethPrimer software21 was used for assay design. Bisulfite-treated DNA was PCR-amplified using unbiased nested primers (SupplementaryTable 2, available as Supplementary data at IJE online), and pyrosequencing was performed to quantitatively measure DNA methylation using the PSQ HS96 (Biotage, Charlotte, NC, USA). Q-CpG methylation software (Biotage, Charlotte, NC, USA) was used to quantify methylation measurements. Mixtures of Whole Genome Amplified (WGA) Human Genomic DNA: Male (Promega, Fitchburg, WI, USA) via REPLI-g Mini Kit (Qiagen, Valencia, CA, USA) and SSIS-treated WGA DNA were used to construct 0, 25, 50, 75 and 100% methylated control standards.

Gene-specific DNA methylation replication in NCS Vanguard samples

The candidate DMR methylation measures at the ANGPT2, PRPF18, FOXD2 and TCEANC2 genes were measured in the independent sample of NCS Vanguard infants, using pyrosequencing methods as described above. Mixed buffy coat and red blood cell aliquots were derived from cord blood collection bags with citrate phosphate dextrose (CPD) anticoagulant. DNA was isolated using the AgencourtGenefind v2 (Beckman Coulter, Pasadena, CA, USA) on the BioMekNNxp laboratory automation workstation according to manufacturer’s instructions. Genomic DNA from each sample was quantified via Picogreen (Invitrogen Life Technologies, Carlsbad, CA, USA). Of 90 cord blood samples, 85 (94%) yielded sufficient DNA for pyrosequencing (200 ng). Genomic DNA from each sample was bisulfite-treated and cleaned using EZ DNA Methylation-Gold Kit (ZymoResearch, Irvine, CA, USA) according to the manufacturer’s instructions. Primers for the additional assays are listed in Supplementary data at IJE online.
Cell type distribution estimation: Genome-scale DNA methylation on the Infinium HumanMethylation450 array

Genome-scale DNA methylation data on NCS Vanguard samples were used to estimate blood cell type distributions per sample. The Infinium HumanMethylation450 assay (Illumina, San Diego, CA, USA) was performed on 147 maternal first and third trimester peripheral blood buffy coat samples and 90 cord blood samples. Agarose gel electrophoresis (1% agarose) was used to check the quality of extracted DNA. Genomic DNA was normalized/aliquoted to 1 μg per sample, and sent to the Johns Hopkins University Center for Inherited Disease Research (CIDR) for bisulfite conversion of genomic DNA with the EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA) and hybridization to the InfiniumBeadChip. As described below, these data were used to estimate cell type proportions per cord blood sample via a cell type prediction algorithm. This algorithm requires reference data from cell type-specific DNA methylation profiles to inform predictions. Cord blood reference samples would be the appropriate input, but no publicly available or proprietary purchasable cord-blood derived or childhood data are currently available. Instead, we used publicly available cell type-specific DNA methylation profiles from adult men.

In addition, we estimated cell type proportions in samples from the THREE study, based on CHARM DNA methylation data. This is an incomplete method, since Infinium probes for which (adult male) cell type reference methylation information is available are not highly overlapping with CHARM probe locations. Cell type proportion estimates obtained in this manner were outside previously reported cord blood cell type ranges and were not biologically plausible. These were not included in downstream analyses.

Gene-specific RNA gene expression

Quantitative real-time PCR (qPCR) was used to measure TCEANC2 mRNA expression levels in a subset of the THREE samples. Fetal cord blood clots were treated with TRIzol (Invitrogen Life Technologies, Carlsbad, CA, USA) and the PureLink RNA Mini Kit (Invitrogen Life Technologies, Carlsbad, CA, USA) was used to isolate RNA, according to the manufacturer’s instructions. cDNA was synthesized using QuantiTect Reverse Transcription Kit (Qiagen, Valencia, CA, USA) with random hexamer priming. Samples with sufficient quantities (n = 42) were run for qPCR in triplicates of technical replicates (Ct SD > 0.25) were excluded (n = 6). Expression levels of TCEANC2 were calculated relative to beta-actin expression. The mercury concentration distributions of 34 THREE participants with final RNA expression results were similar to the discovery methylation population [geometric mean (IQR) methylmercury levels: 0.92 μg/l (0.62–1.34), geometric mean (IQR) total mercury levels: 1.38 μg/l (0.85–1.68)]. Primer sequences used for the qPCR reactions are described in Supplementary Table 3 (available as Supplementary data at IJE online).

Statistical analyses

Study population demographic and exposure characteristics from the discovery THREE and NCS replication sets were compared using t-tests for continuous data and ANOVA tests for categorical variables. In the THREE samples, Spearman correlations between total and methylmercury concentrations were calculated, as were the bivariate relations between total mercury and selected maternal and infant demographic characteristics.

THREE study CHARM 2.0 DNA methylation data were pre-processed, normalized and analysed for regional associations via ‘bumphunting’. In summary, we tested for candidate DMRs associated with total or methylmercury concentration separately by fitting linear models of methylation levels at each probe as a function of log-transformed total or methylmercury concentration levels, adjusted for surrogate variables estimated via SVA. Surrogate variables for downstream adjustment of total mercury and methylmercury analyses were calculated separately and their similarity was tested using Spearman correlation tests. ANOVA tests were performed between measured categorical covariates (array scanner, hybridization date, maternal smoking status, BMI, race, age and offspring mercury, copper, selenium, lead, gestational age, birthweight, and sex) and each surrogate variable. Models did not include adjustment for sex, as we removed sex chromosome probes prior to analysis. Next, estimated regression coefficients from the linear models at each probe within regional probe groups pre-defined in CHARM 2.0 array design were smoothed and filtered by a 99.995th percentile cutoff. Region-level statistics representing the DMR area were calculated and family-wise error rate (FWER) and false discovery rate (FDR) q-values for each DMR were estimated based on 1000 permutations under the null. DMRs that passed the empirical FWER cutoff of 0.2 and FDR q-value cutoff of 0.1 were visualized and targeted for validation and replication. Our primary discovery region-finding analysis involved total mercury levels and our secondary analysis used methylmercury levels.
CHARM DNA methylation measures were validated in a subset of the THREE samples at four genes (TCEANC2, ANGPT2, PRPF18 and FOXD2) using pyrosequencing. A Spearman correlation test of the association between CHARM mean methylation levels across the DMR and mean pyrosequencing methylation across four sites assayed was used to assess concordance between DNA methylation methods.

During the replication phase, we analysed gene-specific NCS DNA methylation data from pyrosequencing data at four genes (TCEANC2, ANGPT2, PRPF18 and FOXD2). We fitted linear models at each CpG, predicting DNA methylation as a function of log-transformed total or methylmercury concentration. In addition, exposure quartiles were calculated and a linear trend test across ordinal methylmercury concentration. We fitted linear models at each CpG, predicting DNA methylation with respect to each CpG. Gene expression values were transformed to log2 scale.

To assess functional implications of methylation differences at the TCEANC2 gene, we calculated Spearman’s rank order correlation between DNA methylation and gene expression with respect to each CpG. Gene expression values were transformed to log2 scale.

We performed several analyses in THREE to address potential confounding by measured covariates including sex, race, maternal age, birthweight, gestational age and concentrations of lead, selenium, copper and EPA + DHA. We tested for association between mercury and potential continuous confounders using Pearson correlation tests of log-transformed mercury measures and between mercury and potential binary confounders using t-tests. Similar tests were performed for association between mean methylation at the TCEANC2 DMR and potential confounders. Finally, we estimated the relationships between average DNA methylation for the DMR and mercury concentration by linear regression, adjusting for each confounder individually in addition to the surrogate variables used in our discovery phase analysis.

There is no publicly available cord blood-derived cell-sorted reference panel to use for cell type estimation similar to what has been applied to adult data.

Population characteristics

Population univariate descriptive statistics for the discovery Baltimore THREE and NCS Initial Vanguard replication samples used for this study are given in Table 1. Distributions of offspring sex and gestational age between the two studies were similar (P-valuerace = 0.9; P-valuegestational age = 0.9). Younger maternal age and lower birthweight were observed in the THREE sample (P-valuerace = 1 × 10−6; P-valuebirth weight = 0.08).

Total and methylmercury concentrations were higher in the cord blood sample from THREE compared with the maternal trimester one blood samples in NCS pregnant women (THREE geometric meanHg = 1.35 μg/l; NCS geometric meanHg = 0.64 μg/l; P-valueHg = 0.0013; and TCEANC2 geometric meanMethylHg = 0.87 μg/l; NCS geometric meanMethylHg = 0.41 μg/l; P-valueMethylHg = 0.029; see Supplementary Table 4, available as Supplementary data at IJE online). Although the distributions differed, both cohorts only contained two samples with methylmercury levels higher than the US Environmental Protection Agency (EPA) reference dose (RfD) for preventing fetal neurotoxicity (5.8 μg/l) and no samples above the health-based benchmark dose (BMD) (85 μg/l) or benchmark dose lower limit 95% confidence interval (BMDL) (58 μg/l). The THREE study had a higher proportion of non-Hispanic Black whereas the NCS Vanguard sample had a higher proportion of non-Hispanic White individuals (P-valuerace = 1 × 10−6).

Within the THREE sample, total and methylmercury concentrations were highly correlated (R = 0.88, P-value = 2 × 10−16). Total mercury level did not vary by infant sex (P-value = 0.47), gestational age (P-value = 0.25), or birthweight (P-value = 0.34). Total mercury was positively associated with maternal age (P-value = 0.0098) and varied by maternal race (P-value = 9.3 × 10−5). Total mercury was elevated among Asians (mean = 2.9 μg/l), intermediate among non-Hispanic Blacks (1.6 μg/l) and lowest among non-Hispanic Whites (1.2 μg/l).
In all, 19 surrogate variables were estimated in the THREE sample for both total and methylmercury analyses. Spearman correlations between surrogate variables for total and methylmercury were high, ranging from 0.986 to 1.000 with a mean of 0.9964 and standard deviation of 0.0049; 14 of the surrogate variables were related to known covariates including hybridization date, array scanner, maternal race, smoking status and BMI (Supplementary Table 5, available as Supplementary data at IJE online).

Genome-scale discovery

In our primary discovery analyses of total mercury and DNA methylation, four DMRs reached our genome-wide significance thresholds (Table 2). At the DMR in the ANGPT2 gene, encoding angiopoietin 2, DNA methylation was positively associated with total mercury levels though the magnitude of methylation change was modest (Supplementary Figure 1a, available as Supplementary data at IJE online). Mean regional methylation at ANGPT2 in the lowest quartile of total mercury concentration (<1.03 μg/l) was 39.1% (SD: 12.1) and the highest quartile of total mercury concentration (>1.99 μg/l) was 41.4% (SD: 9.4). DNA methylation was inversely associated with total mercury levels in a DMR inside the PRPF18 gene, encoding Pre mRNA Processing Factor 18 (Supplementary Figure 1b, available as Supplementary data at IJE online). Lowest quartile mean (SD) was 72.8% (11.7), highest quartile 59.2% (18.6) and ordinal linear

### Table 1. Newborn characteristics of THREE and NCS Vanguard Study samples. Cord blood was used for THREE and whole blood for NCS first-trimester pregnant women

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>THREE, n = 141</th>
<th>NCS, n = 77</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offspring gestational age, median (IQR) (days)</td>
<td>275 (266–282)</td>
<td>275 (270–281)</td>
</tr>
<tr>
<td>Offspring birthweight, median (IQR) (g)</td>
<td>3183 (2794–3566)</td>
<td>3402 (3102–3691)</td>
</tr>
<tr>
<td>Offspring male sex (%)</td>
<td>52</td>
<td>51</td>
</tr>
<tr>
<td>Cord whole blood concentrations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total mercury, median (IQR) (μg/l)</td>
<td>1.4 (1.0–2.0)</td>
<td>–</td>
</tr>
<tr>
<td>Methylmercury, median (IQR) (μg/l)</td>
<td>0.9 (0.6–1.6)</td>
<td>–</td>
</tr>
<tr>
<td>Cord blood serum concentrations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copper, median (IQR) (μg/dl)</td>
<td>39.7 (28.2–53.4)</td>
<td>–</td>
</tr>
<tr>
<td>Selenium, median (IQR) (μg/l)</td>
<td>70.0 (62.0–78.0)</td>
<td>–</td>
</tr>
<tr>
<td>EPA + DHA, median (IQR) (μg/ml)</td>
<td>48.1 (40.8–56.9)</td>
<td>–</td>
</tr>
<tr>
<td>Maternal age, median (IQR) (years)</td>
<td>24 (20–29)</td>
<td>28 (25–33)</td>
</tr>
<tr>
<td>Maternal race (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-Hispanic White</td>
<td>23</td>
<td>75</td>
</tr>
<tr>
<td>non-Hispanic Black</td>
<td>72</td>
<td>6</td>
</tr>
<tr>
<td>Asian</td>
<td>5</td>
<td>–</td>
</tr>
<tr>
<td>Other</td>
<td>–</td>
<td>19</td>
</tr>
<tr>
<td>NCS whole blood concentrations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total mercury, median (IQR) (μg/l)</td>
<td>–</td>
<td>0.6 (0.4–1.1)</td>
</tr>
<tr>
<td>Methylmercury, median (IQR) (μg/l)</td>
<td>–</td>
<td>0.4 (0.2–0.9)</td>
</tr>
<tr>
<td>Selenium, median (IQR) (μg/l)</td>
<td>–</td>
<td>180.5 (165.3–193.3)</td>
</tr>
</tbody>
</table>

IQR, interquartile range.

### Table 2. Candidate differentially methylated regions (DMR) associated with total mercury or methylmercury exposure in the THREE study

<table>
<thead>
<tr>
<th>Exposure variable</th>
<th>Chr</th>
<th>Nearest gene</th>
<th>FWER</th>
<th>q-value</th>
<th>DMR start position</th>
<th>DMR end position</th>
<th>Location relative to gene</th>
<th>Number of probes</th>
<th>Number of CpGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total mercury</td>
<td>8</td>
<td>ANGPT2</td>
<td>0.03</td>
<td>0.045</td>
<td>6418036</td>
<td>6418808</td>
<td>Inside intron</td>
<td>13</td>
<td>42</td>
</tr>
<tr>
<td>Total mercury</td>
<td>10</td>
<td>PRPF18</td>
<td>0.05</td>
<td>0.045</td>
<td>13683914</td>
<td>13684474</td>
<td>Inside intron</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Total mercury</td>
<td>1</td>
<td>FOXD2</td>
<td>0.13</td>
<td>0.078</td>
<td>48059831</td>
<td>48060319</td>
<td>Downstream</td>
<td>8</td>
<td>34</td>
</tr>
<tr>
<td>Total mercury</td>
<td>1</td>
<td>TCEANC2</td>
<td>0.19</td>
<td>0.079</td>
<td>54562102</td>
<td>54562548</td>
<td>Inside exon</td>
<td>7</td>
<td>36</td>
</tr>
<tr>
<td>Methylmercury</td>
<td>1</td>
<td>TCEANC2</td>
<td>0.05</td>
<td>0.088</td>
<td>54562036</td>
<td>54562618</td>
<td>Inside exon</td>
<td>16</td>
<td>44</td>
</tr>
</tbody>
</table>

Chr, chromosome.
trend test \( P_{\text{trend}} = 3.6 \times 10^{-4} \). In addition, an inverse methylation-total mercury association was observed near the FOXD2 gene, which encodes a transcription factor Forkhead Box D2 (Supplementary Figure 1c, available as Supplementary data at IJE online). Lowest quartile mean (SD) was 53.8\% (14.0), highest quartile 43.4\% (17.0) and \( P_{\text{trend}} = 4.6 \times 10^{-3} \). Finally, the exonic DMR associated with TCEANC2, Transcription Elongation Factor A (SII) N-Terminal And Central Domain Containing 2, showed DNA methylation inversely associated with total mercury across a region of 446 base pairs spanning seven probe sets on chromosome 1 (Figure 1a). TCEANC2 regional mean methylation in the lowest quartile of total mercury concentration was 56.7\% (SD: 14.0) and the highest quartile of concentration had 48.0\% (SD: 17.0) methylation (\( P_{\text{trend}} = 9.3 \times 10^{-3} \)).

Our secondary discovery analysis using methylmercury resulted in one methylmercury DMR reaching our genome-wide threshold (Table 2). This methylmercury DMR, located in the TCEANC2 gene, was also identified in the total mercury discovery (Figure 1b). Though the remaining three total mercury DMRs (ANGPT2, PRPF18 and
FOXD2 did not reach genome-wide significance for methylmercury, the magnitudes and directions of methylation change by methylmercury concentration were consistent with those observed for total mercury (Supplementary Figure 1 d–f, available as Supplementary data at IJE online).

Platform validation
Among 136 samples with CHARM and pyrosequencing data, the Spearman correlation coefficient between DNA methylation measures within the TCEANC2 region was 0.82 \((P = 2.2 \times 10^{-16})\). Pyrosequencing results were also correlated with CHARM results for PRPF16 \((\rho = 0.25, P\text{-value} = 0.0040)\) and ANGPT2 \((\rho = 0.18, P\text{-value} = 0.030)\). Pyrosequencing methylation was not correlated with CHARM methylation at FOXD2 \((\rho = 0.08, P\text{-value} = 0.31)\).

Replication
DNA methylation in independent samples from the NCS Vanguard study at four CpGs within the TCEANC2 DMR (red bracket in Figure 1) were inversely associated with both total mercury (Figure 2a) and methylmercury (Figure 2b) concentrations. The direction of association was consistent with the pattern detected during genome-wide discovery pyrosequencing DNA methylation data in NCS Initial Vanguard samples for the ANGPT2, PRPF18 and FOXD2. DMRs were not associated with mercury levels (Supplementary Figure 2, available as Supplementary data at IJE online).

Gene expression
TCEANC2 gene expression was not correlated with paired pyrosequencing DNA methylation data from the four TCEANC2 CpGs (Spearman correlation \(P\text{-value}_{\text{CpG}1} = 0.5; P\text{-value}_{\text{CpG2}} = 0.5; P\text{-value}_{\text{CpG3}} = 0.4; P\text{-value}_{\text{CpG4}} = 0.3; \) Supplementary Figure 3, available as Supplementary data at IJE online). Mean CHARM DNA methylation across the DMR was also not associated with gene expression (Spearman \(\rho = -0.08, P\text{-value} = 0.65\) and expression was not associated with total mercury level (Spearman \(\rho = -0.15, P\text{-value} = 0.38\).

Consideration of confounding
Among the potential demographic and exposure confounders examined, race/ethnicity, blood lead and EPA + DHA fatty acid concentrations showed pairwise association with total and methylmercury \((P\text{-value}_{\text{race-total Hg}} = 1.8 \times 10^{-3}; P\text{-value}_{\text{race-methyl Hg}} = 3.7 \times 10^{-3}; P\text{-value}_{\text{lead-total Hg}} = 0.079; P\text{-value}_{\text{lead-methyl Hg}} = 0.059; P\text{-value}_{\text{EPA+DHA-total Hg}} = 3.6 \times 10^{-3}; P\text{-value}_{\text{EPA+DHA-methyl Hg}} = 5.3 \times 10^{-3}). Of these, race (non-Hispanic White vs other) and EPA + DHA concentration were also associated with mean DNA
methylation at the TCEANC2 DMR (P-value \(= 7.0 \times 10^{-5}\); P-value\(_{\text{EPA+DHA}} = 0.011\)). After adjusting for race or EPA + DHA, the associations between average methylation in this region and total or methylmercury levels were attenuated (Table 3).

DNA methylation at the TCEANC2 DMR does appear to be associated with blood cell type. In the publicly available Reinius et al.\(^{23}\) data containing DNA methylation for six separate blood cell types (B cells, CD4+ T cells, CD8+ T cells, granulocytes, monocytes, natural killer cells and whole blood) from adult blood, methylation among four probe sets at or near the TCEANC2 DMR (chr1: positions 54562102–54562548) varied by cell type (ANOVA: P-value\(_{\text{CpG1}} = 7.7 \times 10^{-3}\); P-value\(_{\text{CpG2}} = 0.046\); P-value\(_{\text{CpG3}} = 0.61\); P-value\(_{\text{CpG4}} = 3.1 \times 10^{-7}\);

**Table 3. Change in DMR effect sizes after adjustment for potential confounders**

<table>
<thead>
<tr>
<th>Model</th>
<th>Total mercury</th>
<th>Methylmercury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Change in percent methylation with 10% increase in exposure (95% CI)</td>
<td>P-value</td>
</tr>
<tr>
<td>Unadjusted</td>
<td>-0.62 (–1.02, –0.23)</td>
<td>3.1 \times 10^{-2}</td>
</tr>
<tr>
<td>Surrogate variables (SV) adjusted(^a)</td>
<td>-0.79 (–1.31, –0.27)</td>
<td>3.7 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + sex of baby</td>
<td>-0.79 (–1.31, –0.26)</td>
<td>3.8 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + maternal age</td>
<td>-0.80 (–1.32, –0.27)</td>
<td>3.5 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + race (non-Hispanic White vs other)</td>
<td>-0.57 (–1.11, –0.037)</td>
<td>3.8 \times 10^{-2}</td>
</tr>
<tr>
<td>SV + birthweight</td>
<td>-0.84 (–1.37, –0.32)</td>
<td>2.2 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + gestational age</td>
<td>-0.83 (–1.35, –0.30)</td>
<td>2.5 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + lead</td>
<td>-0.78 (–1.30, –0.25)</td>
<td>4.8 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + selenium</td>
<td>-0.85 (–1.38, –0.33)</td>
<td>1.9 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + copper</td>
<td>-0.82 (–1.35, –0.29)</td>
<td>2.8 \times 10^{-3}</td>
</tr>
<tr>
<td>SV + EPA+DHA</td>
<td>-0.74 (–1.28, –0.21)</td>
<td>7.5 \times 10^{-3}</td>
</tr>
</tbody>
</table>

DMR methylation calculated as mean percent methylation at TCEANC2. Effect sizes reflect 10% increase in cord blood total or methylmercury concentration.

\(^a\)Model reflecting genome-scale discovery analyses with bump-hunter.

Discussion

We identified a genomic region associated with the TCEANC2 gene where DNA methylation in cord blood is inversely associated with total and methylmercury concentrations. This represents the first genome-scale discovery analysis of DNA methylation and mercury concentration in early life, to the best of our knowledge. The TCEANC2 gene region was identified in the predominantly non-Hispanic Black population from Baltimore, MD, using a genome-wide DNA methylation tiling array and mercury concentrations measured in cord blood. Regional mean methylation in the lowest total mercury quartile was 10% higher than in the highest total mercury quartile group. The association between DNA methylation and mercury at this region was replicated in a predominantly non-Hispanic White US sample using gene-specific bisulfite pyrosequencing and mercury concentrations measured in samples from first-trimester women. Replication across independent populations, using a second method of DNA methylation assessment and a different, but related, exposure window, reduces the likelihood of a false-positive result and increases the generalizability of these results to other populations.

The half-life of methylmercury in blood is estimated to be 70 days\(^{26}\) and 70% or more of blood mercury is in the...
form of methylmercury. Therefore blood mercury concentrations measured during trimester one and cord blood represent non-overlapping windows of susceptibility during the in utero period. Regular exposure to mercury is often assumed to result in a ‘pseudo steady state’ mercury body burden; however, in pregnant women dietary exposures are likely to be intermittent and blood mercury concentrations may vary with time. Cord blood mercury concentrations exceed maternal blood mercury concentrations at birth and due to the rapid placental transfer of lipophilic methylmercury and elevated haematocrit in cord blood, under-scoring the vulnerability of a developing fetus to even low maternal levels of exposure. In a Japanese population, the mean ratio (SD) of total mercury observed in cord blood to maternal blood was 1.6 (0.3). Thus, assuming mercury consumption was at steady state, it would be expected that cord blood concentrations in THREE would be somewhat higher than maternal concentrations in NCS. Mercury processing during pregnancy may vary by racial or ethnic groups, which could be an issue in our study where discovery and replication were performed on different race/ethnicity populations at different points during pregnancy.

Mercury has been associated with altered global and candidate gene-specific DNA methylation in previous laboratory and population studies. Among a population of dental professionals, SEPP1 methylation via bisulfite pyro-sequencing was inversely associated with hair mercury concentrations (representing methyl and inorganic mercury exposure) among males (P-value < 0.05). In our samples, the CHARM platform did not overlap with SEPP1; 15 CHARM probes were within 50 kb of this gene, but none showed association with methylmercury at a P < 0.05 level. Women undergoing in vitro fertilization (IVF) treatments had a positive association between blood mercury and GSTM1/S promoter methylation, measured on the Illumina Golden Gate array (P-value = 0.04). In our study, there were 19 CHARM probes within 5 kb of the GSTM1 gene and 29 probes within 5 kb of the GSTM5 gene. None were associated with methylmercury at P < 0.05. Primary rat neural stem cells in culture exposed to low levels of methylmercury had reduced global methylation. In mouse embryonic stem cells treated with inorganic mercury, DNA hypermethylation at theRnd2 tissue DMR was observed using the COBRA assay. Finally, in our CHARM study, there were 35 probes within 5 kb of the RND2 gene and one was associated with methylmercury exposure at P = 0.0004. The current study expands the genomic coverage of previous work.

The current associations between TCEANC2 methylation and total or methylmercury were robust to several potential confounders including baby’s sex, maternal age, race, gestational age, birthweight, blood lead, selenium and copper concentrations. These results may be confounded by blood cell type composition. Using an adult blood cell reference panel and full-genome Illumina 450 K array data available in the NCS sample, we observed that methylation at the TCEANC2 region varies by estimated blood cell types and that estimated cord blood cell proportions were associated with mercury concentrations. Further, in the NCS sample the association between methylation and mercury in this region was diminished after adjusting for estimated cell type proportions. However, these analyses must be interpreted with caution because it is unknown whether adult reference panels for cell type-specific DNA methylation profiles are appropriate for cord blood. Further work is needed in isolated cell types, using adjustment based on directly measured cord blood cell type proportions or proportions estimated with reference panels from cord blood cell types.

Methylmercury may impair immune function and the developing immune system may be altered by in utero mercury exposure. Thus, one possibility is that the observed association between TCEANC2 DNA methylation and mercury concentration may be related to part to mercury-induced changes in blood cell composition. Adjustment for cell type was not possible in our discovery CHARM analyses; however, we applied an SVA that likely accommodates confounding due to cell type. A large proportion of variance in DNA methylation across the genome can be attributed to cell type, and thus a typical surrogate variable obtained from this approach captures cell-type influences. The results of our study are consistent with either a biological difference in TCEANC2 methylation within blood cell types, or a biological shift in cell type proportions, which may have implications for immune function.

Whereas the TCEANC2 methylation association with mercury was replicated in the independent NCS sample, we did not observe a relationship between TCEANC2 DNA methylation and gene expression in the small number of available cord blood RNA samples (n = 34). There may be gene expression changes with altered methylation at a different developmental time point or in a different tissue or specific cell type. The TCEANC2 protein product is a little studied transcription elongation factor that is also known as C1orf83. The TCEANC2 DMR spans over 400 base pairs within exon 5 and the 3'UTR. The DMR is 42.2 kb upstream of mir4781 and 42.9 kb upstream of the 5'UTR of TMEM59, a gene functionally implicated in Alzheimer’s disease with altered methylation in postmortem Alzheimer’s disease brain tissue. The functional consequences of altered methylation at the TCEANC2 region have not yet been determined; however, infants exposed to methylmercury primarily manifest central nervous system effects and do not have altered haematological
or immunological function. Further studies are needed to investigate any functional consequences of altered methylation at the TCEANC2 gene in the brain and/or if there are brain regions where altered methylation results from methylmercury. An epigenetic change at a specific point in development could knock processes off course (eg neurodevelopmental) in an organ- or time-specific manner; conversely, this mark may be a biomarker of exposure and not aetiological in the expression of the mercury-related central nervous system phenotype.

Blood mercury measurements in this exploratory study were of high quality and reliable, as conducted by an experienced laboratory that makes thousands of similar measurements every year. The THREE and NCS study laboratory measurements were done by the same CDC laboratory that carried out the US National Health and Nutrition Examination Survey (NHANES) measurements for methylmercury and total mercury. There is a considerable overlap between mercury concentrations from our two studies and baseline national levels. Coastal populations in the USA have elevated blood mercury levels relative to the rest of the USA, associated with regional differences in fish consumption. Similarly, we observed higher methylmercury concentrations in the THREE study based out of the coastal city, Baltimore, than the nationwide NCS study. Our results are relevant to mercury methylation blood cell type references.

The current study was limited by cross-sectional biosampling, and repeating the measurements on biosamples collected from the same individuals in the future could demonstrate the persistence or non-persistence of in utero exposure-related DNA methylation marks. In addition, we did not have direct measures or estimates of cell type proportions in the THREE discovery methylation sample because the CHARM methylation platform is not compatible with existing cell-type estimation algorithms. We estimated cell type proportions in the NCS replication sample but, particularly in cord blood, these were limited by: (i) cell type references derived from older, White males; and (ii) admixture of red blood cells in the sample (which are more likely to be nucleated in newborns but not in adults). Future work should generate and use cord-derived DNA methylation blood cell type references.

An additional limitation of our study is that brain rather than blood is the target organ for methylmercury toxicity, though DNA methylation was measured in cord blood. However, in a human population context, cord blood is advantageous as a feasible source of epigenomic information at birth. Primary brain tissue at birth is only available using aborted fetuses or stillbirths (which may have genetic abnormalities), or through study of non-human species (which may respond differently to methylmercury). Certainly confirmatory studies in both human epidemiological samples and via different approaches that would allow evaluation of the central nervous system epigenome are warranted. Long-term follow-up in the NCS and future analyses could provide additional information on DNA methylation persistenec at the TCEANC2 gene and the utility of these marks as a prenatal exposure biomarker in other studies.

Supplementary Data
Supplementary data are available at IJE online.

Funding
This work was supported by the National Institute for Environmental Health Sciences at the National Institutes for Health (ES017646). The National Children’s Study (NCS) Initial Vanguard Centers from which the samples were collected, the NCS biobank team and the NCS repository were supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development at the National Institutes of Health contracts numbers: HHSN27520080033C, HHSN275200503414C, HHSN275200503411C, HHSN275200503416C, HHSN275200503415C, HHSN275200503413C, HHSN275200503410C, HHSN275200503396C, HHSN275201000121U and HHSN27520090010C. Funding for the blood mercury analyses was provided by the National Center for Environmental Health, Centers for Disease Control and Prevention.

Acknowledgements
The manuscript was developed by a writing team identified by the National Children’s Study Publications Committee for the purpose of timely sharing of centrally collected NCS data. We acknowledge the contributions of the following Vanguard Centers and principal investigators: Children’s Hospital of Philadelphia, Jennifer Culhane; Mt. Sinai Medical School, Phil Landrigan; South Dakota State University, Bonny Specker; University of California at Irvine, James Swanson and Dean Baker; University of North Carolina at Chapel Hill, Barbara Entwisle and Nancy Dole; University of Utah School of Medicine, Ed Clark; University of Wisconsin, Maureen Durkin.

Conflict of interest: None declared.

References