**SUPPLEMENTAL METHODS**

**Integrative Exposomic, Genomic, Epigenomic Analyses of Human Placenta Samples Links Understudied Chemicals to Preeclampsia**

Alex Chaoa,\*, Jarod Grossmanb, Celeste Carberryc,d, Yunjia Laic, Antony J. Williamsa, Jeffrey M. Minuccie, S. Thomas Puruckerf, John Szilagyic,d, Kun Luc,d,g,, Kim Boggessh, Rebecca C. Fryc,d,g, Jon R. Sobusa,Julia E. Rager c,d,g,\*

**Affiliations**

aU.S. Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure, Chemical Characterization and Exposure Division, Research Triangle Park, NC, USA

bAgilent Technologies, Santa Clara, CA, USA

cDepartment of Environmental Sciences and Engineering, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

dThe Institute for Environmental Health Solutions, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

eU.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Public Health and Environmental Systems Division, Research Triangle Park, NC, USA

fU.S. Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division, Research Triangle Park, NC, USA

gCurriculum in Toxicology and Environmental Medicine, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

hDepartment of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

\*Co-corresponding authors at

AC: 109 TW Alexander Dr., RTP, NC 27711, Mail Code: E205-04; chao.alex@epa.gov

JR: 135 Dauer Drive, CB 7431, Chapel Hill, NC, 27599; jrager@unc.edu

**SUPPLEMENTAL METHODS TABLE OF CONTENTS**

**Section 1** - Supplemental Methods for High-Resolution Mass Spectrometry Analysis using Non-Targeted Methods – Page 3

**Section 2** - Supplemental Methods for High-Resolution Mass Spectrometry Analysis using Targeted Methods – Page 7

**Section 3** - Supplemental Methods for Genomic and Epigenomic Profiling of Placenta Samples – Page 8

**Section 4** - Feature Intensity Results for Tracer Compounds – Page 9

**Section 5** - References for Supplemental Methods – Page 13

**Section 1 - Supplemental Methods for High-Resolution Mass Spectrometry Analysis using Non-Targeted Methods**

Placenta samples were analyzed using an Agilent 1290 Infinity II ultra-high-performance liquid chromatograph (UHPLC) coupled to an Agilent 6546 high-resolution quadrupole time-of-flight (Q-TOF) mass spectrometer with a dual AJS electrospray ionization source. Source and method parameters were as follows: drying gas temperature = 300°C; drying gas flow = 9 L/min; nebulizer gas = 35 psi; sheath gas temperature = 350°C; sheath gas flow = 11 L/min; Vcap = 3500V (3000V in negative mode); nozzle voltage = 0V (2000V in negative mode); and fragmentor = 100V. Column parameters: An Agilent Zorbax Eclipse Plus C8 (2.1 x 100 mm, 1.8 µm) UHPLC column was used for compound separation (part number: 959758-906) and the column temperature was set to 40°C. Method parameters: Mobile phases consisted of (A) deionized water (Millipore MilliQ Water System) with 0.1% formic acid (Fisher) and (B) 90:10 acetonitrile (Alfa Aesar):deionized water with 0.1% formic acid for LCMS analyses run in positive ionization mode. For analyses run in negative ionization mode, mobile phases were identical with the addition of 0.5 mM ammonium fluoride (Fluka) to both mobile phases. The flow rate was set at 0.5 mL/min, and mobile phases were ramped with the following gradient: 0-1 min = 5% B; 1-7 min 5-80% B ramp; 7-9 min = 80-90% B ramp; 9-10 min = 90-100% B ramp; 10-12 min = 100% B; and 12-13.5 min = 5% B.

Each sample extract was first analyzed in full scan mode (MS1) with triplicate injections (2 µL) in each ionization mode. A fully randomized run sequence was used, which included replicates of solvent blanks, method blanks, and pooled QC samples. For MS1 analysis, the scan range was 100-1000 m/z with a scan speed of 2 spectra/sec. A reference solution consisting of purine (DTXSID5074470), hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazene (DTXSID90880494), and trifluoroacetic acid (DTXSID9041578) was infused into the source during each run to auto-correct mass drift. Data-dependent acquisition MS2 (auto MSMS) was also performed on each sample extract in both positive and negative ionization modes, cycling between collision energies of 10, 20, and 40 eV. MS2 data were acquired with the following settings: scan range 40-1000 m/z; 6 max precursors per cycle; minimum threshold 3000 counts; scan speed 5 spectra/sec, medium isolation width (4 m/z); active exclusion enabled after 3 MS2 spectra; exclusion release time 0.2 minutes. MS2 exclusion lists were used to exclude ions corresponding to compounds in the reference solution. Raw data files from LC-MS runs have been submitted to Metabolomics Workbench (Study ID: ST002151, DOI: 10.21228/M8DD7D, mwTab filename: alexchao\_20220422\_055158\_mwtab.txt, study title: Integrative Exposomic, Transcriptomic, Epigenomic Analyses of Human Placental Samples Links Understudied Chemicals to Preeclampsia).

A solution of eleven isotopically labeled tracer compounds was spiked into each blank, QC, and study sample to assess mass accuracy, retention time drift, and measurement precision (using the percent coefficient of variation [CV] of feature intensities across triplicate injections) over the course of the run. The following table lists the compounds contained within the tracer solution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Chemical Name** | **DTXSIDe** | **Ionization Mode** | **Mass Error (ppm)** | **Precision (max %CV)** |
| 13C6-Methyl parabena | DTXSID30894090 | ESI- | 0.00 | 11% |
| 13C4-Perfluorooctanoic acidb  | DTXSID70892999 | ESI- | 2.63 | 9% |
| 13C4 15N2-Fipronila | DTXSID50894092 | ESI- | 1.36 | 8% |
| 13C4 15N2-Fipronil sulfonea | DTXSID10894093 | ESI- | 3.06 | 11% |
| 13C4-Perfluorooctanesulfonic acidb  | DTXSID80894101 | ESI- | 2.98 | 6% |
| 13C2-Perfluorodecanoic acidb | DTXSID20894100 | ESI- | 2.13 | 12% |
| 13C3-Atrazinea | DTXSID60894088 | ESI+ | 0.83 | 2% |
| D10-Carbamazepinec | DTXSID30497060 | ESI+ | 1.14 | 5% |
| D5-Diphenhydramined | DTXSID601034131 | ESI+ | 1.19 | 4% |
| D3-Thiamethoxamd | DTXSID60746816 | ESI+ | 0.27 | 8% |
| D4-Pyriproxyfend | DTXSID20894089 | ESI+ | 1.81 | 4% |

a Cambridge Isotope Laboratories, Inc. (Tewksbury, MA, USA)

b Wellington Laboratories Inc. (Guelph, Ontario, Canada)

c Sigma-Aldrich (St. Louis, MO, USA)

d CDN Isotopes Inc. (Pointe-Claire, Quebec, Canada)

e DSSTox Substance Identifier

For QA/QC purposes, one study sample was prepared in duplicate (Sample A and B) and further split into a “Pre” and “Post” sample, with the “Pre” sample spiked with tracer solution prior to extraction, and the “Post” sample spiked following extraction (as was performed for all other study samples). These four preparations (“PreA”, “PreB”, “PostA”, “PostB”) were each analyzed in triplicate, yielding twelve total injections. Discrepancies in feature intensities for each tracer between “Pre” and “Post” samples reflect levels of experimental recovery (i.e., extraction + filtration). Only data from the “PostA” sample injections were included in statistical analyses involving all other study samples. Graphs of feature intensities for tracers across all samples are shown in Supplemental Methods Section 4.

Including triplicate injections of one method blank, all QC samples, and all study samples, a total of 123 sample injections were evaluated as part of this study. All eleven tracers were detected in all 123 sample injections. Feature intensities were extremely stable for all eleven tracers across the sample injections (not considering six “PreA” and “PreB” injections, where intensities were noticeably diminished), with ESI+ mode tracer CV values ≤ 0.08 and ESI- mode values ≤ 0.12 (Tables S7 and S8). Given this stability, no injection-specific corrections were considered for this data set. Mass error for the eleven tracer compounds was ≤ 3ppm and retention time drift was ≤ 0.1 min (Tables S7 and S8).

Agilent Profinder v10 was used for peak picking and alignment of NTA data. As part of the peak picking process, peak areas were integrated to generate intensity values for each peak. An artifact of the recursive peak picking process within Profinder is the occasional generation of duplicate features. Following peak picking, the NTA WebApp compares features for duplicate values to identify superfluous features for removal. The following tables list settings used for Profinder processing:

|  |  |
| --- | --- |
| **Step 1** | **MFE Extraction Parameters** |
| Peak Height | 1000 |
| Ions (+) | H+ |
| Ions (-) | H- |
| Isotope Model | Common organic molecules |
| Charge state |  1-2 |
| **Step 2** | **MFE Compound Filters** |
| Compound Ion Count | Include all |
| Compounds with indeterminate neutral mass | Exclude |
| **Step 3** | **Compound Binning and Alignment** |
| RT Tolerance |  +/- (0.00% + 0.03 min) |
| Mass Tolerance |  +/- (15.00 ppm + 2.00 mDa) |
| **Step 4** | **MFE Post-Processing Filters** |
| Score (MFE) | 70 |
| Minimum filter matches | 100% of files in at least one sample group (samples grouped as replicates and also as types) |
| **Step 5** | **Find by Ion Matching Tolerances and Scoring** |
| Match Tolerance (Masses) |  +/- 10.00 ppm |
| Match Tolerance (Retention times) |  +/- 0.03 minutes |
| Expansion for chromatographic extraction (Possible m/z) | Symmetric (ppm) +/- 35 |
| Limit EIC extraction range | Yes |
| Expansion for chromatographic extraction (Expected retention time) | Symmetric +/- 0.75 minutes |
| Mass score | 100 |
| Isotope abundance score | 60 |
| Isotope spacing score | 50 |
| Retention time score | 0 |
| Expected data variation (MS mass) | 2.0 mDA + 5.6 ppm |
| Expected data variation (MS isotope abundance) | 7.50% |
| Expected data variation (Retention time) | 0.115 min |
| Warn if a score is  | <75.00 |
| Warn if the (unobserved) second ion's abundance is expected to be | >1000.00 |
| **Step 6** | **Find by Ion EIC Peak Integration and Filtering** |
| Integration algorithm | Agile 2 |
| Smoothing | Gaussian (Function width 9 points) (Gaussian width 5 points) |
| Absolute height EIC filter | 3000 counts |
| **Step 7** | **Find by Ion Spectrum Extraction and Centroiding** |
| Spectra to include | Average scans > 10% of peak height |
| Exclude if above | 20% of saturation in the m/z ranges used in the chromatogram, never return an empty spectrum |
| Peak spectrum background | none |
| Centroiding Peak location Maximum spike width | 2 |
| Centroiding Peak location Required valley | 0.7 |
| **Step 8** | **Find by Ion Post-Processing Filters** |
| Score (Tgt) | 50 |
| Minimum filter matches | 100% of files in at least one sample group (samples grouped as replicates and also as types) |

Agilent Mass Profiler Professional (MPP) v8 was used for assigning chemical formula to aligned molecular features. The following tables list settings used for MPP processing:

|  |  |
| --- | --- |
| **Step 1** | **Data Import Filtering** |
| Retention Time | Use all available data |
| Mass | Use all available data |
| Minimum number of ions | 2 |
| Charge states | All charge states permitted |
| **Step 2** | **Alignment Parameters** |
| RT Window | 0.0% + 0.01 min |
| Mass Window | 0.01 ppm + 0 mDa |
| **Step 3** | **Compound Identification** |
| Compound identification methods |  Database search |
| Mass Tolerance |  +/- (10 ppm + 2.00 mDa) |
| Database | DSSTox MS-Ready Mapping File (downloaded on 2/21/2019) |
| Positive Ions | +H  |
| Negative Ions | -H |
| Charge state range | 1-2 |
| Maximum number of peaks to search when peaks are not specified graphically | 5 |
| Search Results: Limit to the best | 10 hits |
| **Step 4** | **Compound Identification: Scoring** |
| Mass score | 100.00 |
| Isotope abundance score | 60.00 |
| Isotope spacing score | 50.00 |
| Retention time score | 100.00 |
| Expected data variation: MS mass | 2.0 mDa + 5.6 ppm |
| MS isotope abundance | 7.5% |
| MS/MS mass | 5.0 mDa + 7.5 ppm |
| Retention time | 0.115 min |

**Section 2 - Supplemental Methods for High-Resolution Mass Spectrometry Analysis using Targeted Methods**

Placenta samples and confirmation standard mixtures were analyzed using a Thermo Vanquish UHPLC coupled to a Thermo Q-Exactive Orbitrap mass spectrometer. Source and method parameters were as follows: sheath gas = 60 L/min; aux gas = 10 L/min; sweep gas = 1 L/min; spray voltage = 3.75 kV; spray current = 2.5 µA; capillary temp = 325ºC; and aux gas heater temp = 400ºC. Column parameters and mobile phases were as described for non-targeted methods. The flow rate was set at 0.5 mL/min, and mobile phases were ramped with the following gradient: 0-1 min = 5% B; 1-7 min 5-80% B ramp; 7-9 min = 80-90% B ramp; 9-10 min = 90-100% B ramp; 10-11 min = 100% B; 11-12 min = 100-5% ramp; and 12-13.5 min = 5% B.

Each sample extract and standard mixture were analyzed in full scan mode (MS1) and hybrid mode (MS1/MS2, i.e., full scan alternating with PRM) with injection volumes of 2 µL. For MS1 analysis the following parameters were used: AGC target = 3e6; maxIT = 225 ms; and mass resolution (FWHM) = 70,000. For hybrid mode, MS1 full scan was same as above; for MS2 PRM, AGC target = 2e5, maxIT = 100 ms, isolation window = 1.2 m/z, resolution = 17,500, stepped NCE = 10, 50, 100. Targeted m/z values were determined for each compound by considering [M+H]+ and [M-H]- ion species for positive and negative ionization mode, respectively.

**Section 3 - Supplemental Methods for Genomic and Epigenomic Profiling of Placenta Samples**

Nucleic acids were extracted from the same placentas as used for chemical analyses as previously described (Brooks et al. 2016; Martin et al. 2015). In brief, a subsection (approximately 0.2g) was sliced from each frozen tissue sample, placed on dry ice using a sterile dermal curette, and washed in sterile 1x PBS to remove any potential blood contamination. Samples were then immediately homogenized in Buffer RLT containing b-mercaptoethanol using a TissueRuptor (Qiagen, Valencia, CA). Nucleic acids were extracted from the resulting homogenate using the AllPrep DNA/RNA/miRNA Universal Kit according to the manufacturer’s instructions (Qiagen), which yielded RNA sequences 18 nucleotides and greater in length as well as DNA from the sample tissue samples. Extracted DNA and RNA were quantified with a Nanodrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA) and RNA integrity verified with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA).

Genome-wide gene expression profiles were evaluated by labeling and hybridizing RNA samples to the Affymetrix GeneChipTM Human Gene 2.0 ST array as previously described (Brooks et al. 2016; Martin et al. 2015). This microarray assesses the relative expression levels of >25,000 mRNAs using 53,617 probesets. Resulting gene expression data were normalized by robust multi-chip average (Irizarry et al. 2003) and log2-transformed prior to statistical evaluation. Genome-wide miRNA expression profiles were evaluated by labeling and hybridizing RNA samples to the Agilent Human miRNA Microarray (v16) as previously described (Brooks et al. 2016). This microarray assesses the relative expression levels of 1347 miRNAs using 56,044 probesets. Resulting miRNA expression data were collapsed according to individual miRNA using the median signal, normalized by quantiles, and log2-transformed prior to statistical evaluation, as done previously (Brooks et al. 2016; Rager et al. 2014). Differential RNA expression was statistically defined using a false discovery rate (FDR) corrected q ≤ 0.15.

Genome-wide CpG methylation profiles were evaluated using the Infinium HumanMethylation450 BeadChip array (Illumina, Inc., San Diego, CA) as previously described (Martin et al. 2015). This array assesses the CpG methylation levels of 486,428 sites across the DNA at single nucleotide resolution. Prior to array hybridization, DNA was first bisulfite-converted using the EZ DNA methylation kit (Zymo Research, Irvine, CA). Resulting methylation array data were produced as β values, where β = intensity of the methylated allele (M) / (intensity of the unmethylated allele (U) + intensity of the methylated allele (M) + 100). Data were pre-processed in accordance with an established 450k Chip Analysis Methylation Pipeline (Morris et al. 2014). Probes within sites of known single nucleotide polymorphisms were removed, as were probes with detection p-values > 0.01 to control for probe quality, leaving 390,452 probes that were then quantile normalized prior to statistical evaluation. The resulting 390,452 quantile normalized probes were assessed for differential methylation in relation to preeclampsia case status using an Analysis of Covariance (ANCOVA) model controlling for gestational age, maternal age, and race. Differential DNA methylation levels were statistically defined as: a p-value <0.01 for associations with case status with a false discovery rate (FDR) corrected q-value < 0.1.

**Section 4 – Performance Metrics for HRMS NTA**

Tracer compounds spiked into all samples were used to track and assess instrument performance during acquisition of NTA data. Relevant metrics pertaining to each compound are shown here as feature intensity results for tracer compounds. Also included are additional performance metrics in **Supplemental Tables S7** and **S8**.

****

****

****

****

****

****

****

****

****

****

****

**Section 5 – References for Supplemental Methods**

Brooks SA, Martin E, Smeester L, Grace MR, Boggess K, Fry RC. 2016. Mirnas as common regulators of the transforming growth factor (tgf)-beta pathway in the preeclamptic placenta and cadmium-treated trophoblasts: Links between the environment, the epigenome and preeclampsia. Food Chem Toxicol 98:50-57.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. Summaries of affymetrix genechip probe level data. Nucleic Acids Res 31:e15.

Martin E, Ray PD, Smeester L, Grace MR, Boggess K, Fry RC. 2015. Epigenetics and preeclampsia: Defining functional epimutations in the preeclamptic placenta related to the tgf-beta pathway. PLoS One 10:e0141294.

Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. 2014. Champ: 450k chip analysis methylation pipeline. Bioinformatics 30:428-430.

Rager JE, Bailey KA, Smeester L, Miller SK, Parker JS, Laine JE, et al. 2014. Prenatal arsenic exposure and the epigenome: Altered micrornas associated with innate and adaptive immune signaling in newborn cord blood. Environ Mol Mutagen 55:196-208.