Supplementary Methods:

Amniotic fluid samples were initially processed at Metabolon, Inc using their standard platform described below.

Amniotic Fluid sample accessioning and preparation: Following receipt, samples were inventoried and immediately stored at -80°C. Each sample received was accessioned into the Metabolon LIMS system and was assigned by the LIMS a unique identifier Samples were prepared using the automated MicroLab STAR® system from Hamilton Company. A recovery standard was added prior to the first step in the extraction process for QC purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol (5:1 methanol: AF by volume) under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: one for analysis by UPLC-MS/MS with positive ion mode electrospray ionization, one for analysis by UPLC-MS/MS with negative ion mode electrospray ionization, one for LC polar platform, one for analysis by GC-MS, and one sample was reserved for backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. For LC, the samples were stored overnight under nitrogen before preparation for analysis. For GC, each sample was dried under vacuum overnight before preparation for analysis.

QA/QC: Several types of controls were analyzed in concert with the experimental samples: a pooled matrix sample generated by taking a small volume of each experimental sample (or alternatively, use of a pool of well-characterized human plasma) served as a technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of QC standards that were carefully chosen not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowed instrument performance monitoring and aided chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e., non-instrument standards) present in 100% of the pooled matrix samples. Experimental samples were randomized across the platform run with QC samples spaced evenly among the injections.

Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS): The LC/MS portion of the platform was based on a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The LC/MS gradients were described in a prior publication (Evans AM et al, "Integrated, Nontargeted Ultrahigh Performance Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry Platform for the Identification and Relative Quantification of the Small-Molecule Complement of Biological Systems" Anna. Chem 2009, 81, 6656-6667). The sample extract was dried then reconstituted in acidic or basic LC-compatible solvents, each of which contained 8 or more injection standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion optimized conditions and the other using basic negative ion optimized conditions in two independent injections using separate dedicated columns (Waters UPLC BEH C18-2.1x100 mm, 1.7 μm). Extracts reconstituted in acidic conditions were gradient eluted from a C18 column using water and methanol containing 0.1% formic acid. The basic extracts were similarly eluted from C18 using methanol and water, however with 6.5mM Ammonium Bicarbonate. The third aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 μm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate. The MS analysis alternated between MS and data-dependent MS2 scans using dynamic exclusion, and the scan range was from 80-1000 m/z. Raw data files are archived and extracted as described below.

Gas Chromatography-Mass Spectroscopy (GC-MS): The samples destined for analysis by GC-MS were dried under vacuum for a minimum of 18 h prior to being derivatized under dried nitrogen using bistrimethyl-silyltrifluoroacetamide. Derivatized samples were separated on a 5% diphenyl / 95% dimethyl polysiloxane fused silica column (20 m x 0.18 mm ID; 0.18 um film thickness) with helium as carrier gas and a temperature ramp from 60° to 340°C in a 17.5 min period. Samples were analyzed on a Thermo-Finnigan Trace DSQ fast-scanning single-quadrupole mass spectrometer using electron impact ionization (EI) and operated at unit mass

resolving power. The scan range was from 50–750 m/z. Raw data files are archived and extracted as described below.

Bioinformatics: The informatics system consisted of four major components, the Laboratory Information Management System (LIMS), the data extraction and peak-identification software, data processing tools for QC and compound identification, and a collection of information interpretation and visualization tools for use by data analysts. The hardware and software foundations for these informatics components were the LAN backbone, and a database server running Oracle 10.2.0.1 Enterprise Edition.

LIMS: The purpose of the Metabolon LIMS system was to enable fully auditable laboratory automation through a secure, easy to use, and highly specialized system. The scope of the Metabolon LIMS system encompasses sample accessioning, sample preparation and instrumental analysis and reporting and advanced data analysis. All of the subsequent software systems are grounded in the LIMS data structures. It has been modified to leverage and interface with the in-house information extraction and data visualization systems, as well as third party instrumentation and data analysis software.

Data Extraction and Compound Identification: Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contain the retention time/index (RI), mass to charge ratio (*m/z*), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 0.005 amu, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. More than 3300 commercially available purified standard compounds have been acquired and registered into LIMS for distribution to both the LC-MS and GC-MS platforms for determination of their analytical characteristics. Additional mass spectral entries have been created for structurally unnamed biochemicals, which have been identified by virtue of their recurrent nature (both chromatographic and mass spectral). These compounds have the potential to be identified by future acquisition of a matching purified standard or by classical structural analysis.

Curation: A variety of curation procedures were carried out to ensure that a high quality data set was made available for statistical analysis and data interpretation. The QC and curation processes were designed to ensure accurate and consistent identification of true chemical entities, and to remove those representing system artifacts, mis-assignments, and background noise. Metabolon data analysts use proprietary visualization and interpretation software to confirm the consistency of peak identification among the various samples. Library matches for each compound were checked for each sample and corrected if necessary. Metabolite quantification was performed by measuring area-under-the-curve for each peak.

Statistical Analysis: For each of the 459 named biochemical detected, the data was log transformation and missing values, if detected, were imputed with the minimum observed value for each compound, Two way ANOVA contrasts were used to identify biochemicals that differed significantly between experimental groups looking at both GDM status and sex of the offspring. A p value < 0.05 was considered significant. An estimate of the false discovery rate (*q*-value) was calculated to take into account the multiple comparisons that normally occur in metabolomic-based studies and is described as a *q*-value. The FDR for a given set of compounds can be estimated using the q-value (Storey J and Tibshirani R. (2003) Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. USA 100: 9440-9445; PMID: 12883005).

Random Forest: Random forest is a supervised classification technique based on an ensemble of decision trees (see Breiman L. (2001) Random Forests. Machine Learning. 45: 5-32; http://link.springer.com/article/10.1023%2FA%3A1010933404324). Random Forest analysis was generated using Array Studio with a script written in R. For a given decision tree, a random subset of the data with identifying true class information is selected to build the tree ("bootstrap sample" or "training set"), and then the

remaining data, the "out-of-bag" (OOB) variables, are passed down the tree to obtain a class prediction for each sample. This process is repeated thousands of times to produce the forest. The final classification of each sample is determined by computing the class prediction frequency ("votes") for the OOB variables over the whole forest. To determine which variables (biochemicals) make the largest contribution to the classification, a "variable importance" measure is computed. We use the "Mean Decrease Accuracy" (MDA) as this metric. The MDA is determined by randomly permuting a variable, running the observed values through the trees, and then reassessing the prediction accuracy. If a variable is not important, then this procedure will have little change in the accuracy of the class prediction (permuting random noise will give random noise). By contrast, if a variable is important to the classification, the prediction accuracy will drop after such a permutation, which we record as the MDA. Thus, the random forest analysis provides an "importance" rank ordering of biochemicals; we typically output the top 30 biochemicals in the list as potentially worthy of further investigation.

Principal Components Analysis (PCA): PCA is an unsupervised analysis that reduces the dimension of the data. Each principal component is a linear combination of every metabolite and the principal components are uncorrelated. The first principal component is computed by determining the coefficients of the metabolites that maximizes the variance of the linear combination.   The second component finds the coefficients that maximize the variance with the condition that the second component is orthogonal to the first.   The third component is orthogonal to the first two components and so on. The total variance is defined as the sum of the variances of the predicted values of each component (the variance is the square of the standard deviation), and for each component, the proportion of the total variance is computed.   For example, if the standard deviation of the predicted values of the first principal component is 0.4 and the total variance = 1, then 100*0.4*0.4/1 = 16% of the total variance is explained by the first component. Since this is an unsupervised method, the main components may be unrelated to the treatment groups, and the "separation" does not give an estimate of the true predictive ability. PCA was generated using Array Studio with a script written in R.