Detection and Confirmation of Serum Lipid Biomarkers for Preeclampsia Using Direct Infusion Mass Spectrometry

Swati Anand, 1Sydney Young, 2, 3 Sean M. Esplin, 4Benjamin Peaden, 4H. Dennis Tolley, 2, 3T. Flint Porter, 2, 3Michael W. Varner, 5Mary E. D’Alton, 1Bruce Jackson, 1Steven W. Graves

1Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT; 2Maternal Fetal Medicine, Intermountain Healthcare, Salt Lake City, UT; 3Department of Obstetrics and Gynecology, University of Utah Health Sciences Center, Salt Lake City, UT; 4Department of Statistics, Brigham Young University, Provo, UT; 5Department of Obstetrics and Gynecology, Columbia University School of Medicine, New York, NY.

Corresponding Author:
Dr. Steven W. Graves
Department of Chemistry and Biochemistry
BNSN C-212
Brigham Young University
Provo, UT
USA 84602
Tel # 801-422-2148
FAX # 801-422-0153
email: swgraves@chem.byu.edu

Running Title: Predictive serum lipidomic biomarkers for preeclampsia

Abbreviations: MS, mass spectrometry; ROS, reactive oxygen species; TOF, time-of-flight; LC/MS, liquid chromatography/mass spectrometry; ESI, electrospray ionization; TIC, total ion chromatogram; MCA; multi-channel analyzer; ROC, receiver operator curves; PC, phosphatidylcholines; AUC, area under the curve.
ABSTRACT

Despite substantial research, the early diagnosis of preeclampsia remains elusive. Lipids are now recognized to be involved in regulation and pathophysiology of some disease. Shotgun lipidomic studies were undertaken to determine if serum lipid biomarkers exist that predict preeclampsia later in the same pregnancy. A discovery study was performed using sera collected at 12-14 weeks pregnancy from 27 controls with uncomplicated pregnancies and 29 cases that later developed preeclampsia. Lipids were extracted and analyzed by direct infusion into a time-of-flight mass spectrometer. MS signals, demonstrating apparent differences were selected, their abundances determined and statistical differences tested. Statistically significant lipid markers were reevaluated in a second confirmatory study having 43 controls and 37 preeclampsia cases. Multi-marker combinations were developed using those lipid biomarkers confirmed in the second study. The initial study detected 45 potential preeclampsia markers. Of these, 23 markers continued to be statistically significant in the second confirmatory set. Most of these markers representing several lipid classes were chemically characterized typically providing lipid class and potential molecular components using tandem MS. Several multi-marker panels with AUC>0.85 and high predictive values were developed. Developed panels of serum lipidomic biomarkers appear able to identify most women at risk for preeclampsia in a given pregnancy at 12-14 weeks gestation.

Keywords: Preeclampsia, Diagnosis, Biomarkers, Lipidomics, Pregnancy, Mass spectrometer and Lipids
INTRODUCTION

Preeclampsia is a potentially life-threatening disorder of pregnancy characterized by new-onset hypertension and proteinuria after 20 weeks gestation. It constitutes a leading cause of maternal and perinatal mortality and morbidity (1-2). Estimates are that as many as 75,000 women worldwide die each year from preeclampsia (3). Treatment options are very limited and frequently require termination of the pregnancy, regardless of gestational age, accounting for ~20% of all preterm births (4). Furthermore, infants born to preeclamptic mothers may be at an increased risk of hypertension, heart disease and diabetes beyond their being premature (5).

While there are known risk factors for preeclampsia, it is not yet possible to precisely distinguish pregnancies destined to develop preeclampsia from those that will not. This has made prospective clinical studies large and expensive. Also, it has to some degree limited studies of early changes that may lead to preeclampsia. In the absence of a complete animal model of preeclampsia, the actual cause or causes of this disorder remains unknown. The pathogenesis is acknowledged to be complex and although incompletely understood, it is generally believed that one or more very early events in the pregnancy contribute to preeclampsia. One long held theory to explain this disease involves an incomplete remodeling of maternal spiral arteries by invasive extravillous placental trophoblast cells resulting in inadequate perfusion of the fetal-placental unit with attendant ischemia (6-7). Abnormal waveform patterns and an increased pulsatility index observed in uterine Doppler ultrasound studies support the concept of an underperfused fetus prior to clinically apparent preeclampsia (6, 7). There is also evidence to suggest that biochemical changes may be seen in women who, later in the same pregnancy, develop this disease. Several other biochemical changes appear to precede by a few weeks clinically evident preeclampsia, including hypoxia-reoxygenation (8), abnormal expression of angiogenic and anti-angiogenic factors in the maternal circulation and endothelial dysfunction with a pro-inflammatory response (9-10). Collectively, these data suggest that there are biochemical abnormalities that occur prior to clinically evident preeclampsia signs and symptoms. However, to date there are still no accepted, predictive biomarkers for this disease (11-13).
despite some initial promise for a number of proposed candidates. Also currently, the specific causes (as opposed to consequences) of preeclampsia are still debated and have yet to be established.

It is now possible to survey hundreds to thousands of molecules in biological specimens in an unbiased way using mass spectrometry (MS). Proteomics is by far and away the most common of these approaches and has been employed to study several diseases, but its use to explore diagnostic or predictive biomarkers in serum is difficult due to ~30 highly abundant serum proteins that mask the vast majority of lower abundance species in serum due to ion suppression. Other methods have targeted peptides and lipids, although much less frequently in serum or plasma.

Lipids are increasingly recognized as having important biological roles or representing important biochemical correlates of clinical changes. A wide variety of human diseases are associated with aberrant lipid metabolism including Alzheimer’s disease, diabetes and atherosclerosis (14-16). Alterations in lipids may represent by-products of underlying pathophysiology but could also represent primary disease mediators. For example, arachidonic acid is a precursor for eicosanoids that have a significant role in inflammatory processes. There is evidence of changes in downstream products of arachidonic acid in preeclampsia (17). As another example, oxidized lipid species can reflect increased reactive oxygen species (ROS) which are produced as a part of several diseases (18-20). Therefore, there is ample reason to study lipid profiles in pregnant women at an early stage seeking predictive lipidomic biomarkers to identify patients at significant risk for preeclampsia. Consequently, a ‘global’ serum lipidomic approach, involving lipid extraction followed by direct injection, time-of-flight mass spectrometry was used to identify and chemically characterize predictive serum preeclamptic lipid biomarkers. We hypothesized that this approach would find individual lipids and sets of lipids that would allow for the prediction of a substantial portion of women who would later develop this disease in the same pregnancy and that the changes would provide insights into the mechanisms early in the development of preeclampsia.
METHODS

Study population

Serum specimens used for both the discovery and confirmatory studies were obtained from the Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT and from Intermountain Health Care (Intermountain), Murray, UT. All samples were banked sera, obtained from a previously completed clinical study, “First- And Second-Trimester Evaluation of Risk” (21). Institutional Review Board (IRB) approval for these studies was obtained from the University of Utah, Intermountain and Brigham Young University prior to our initiating these experiments. Sera had been collected from pregnant women at 12-14 weeks gestation followed through the completion of their pregnancies. Specimens were analyzed without clinical identifiers.

A discovery cohort involved sera from 27 controls, having term, uncomplicated pregnancies and 29 cases, who developed preeclampsia later during the index pregnancy. The second confirmatory study of the promising potential lipid biomarkers from the discovery cohort involved serum from 37 cases and 43 controls, also collected at 12-14 weeks gestation. Demographic characteristics are summarized in Tables 1a and 1b.

Materials

Glycerophosphocholine lipid standards: PC (14:0/16:0) and PC (18:0/18:2) were purchased from Avanti Lipids (Alabaster, AL, USA). A triacylglycerol standard: TG (16:0/18:0/16:0) was purchased from Sigma Aldrich (St. Louis, MO, USA).

Sample preparation

All specimens had been stored at -80°C until our receiving them on dry ice and were maintained at -80°C before and after sample processing. Serum lipid extraction efficiency of five different organic extraction solutions were performed as follows: for a 200 µL sample 1) 2.5 mL of 4:1 chloroform:methanol; 2) 2.5 mL of 4:1 methyl-tert-butyl ether:methanol; 3) 1.8 mL of 3:2 hexane:isopropanol; 4) 2.3 mL of...
3.6:1 hexane:methanol); or 5) 2.0 mL of 3:1 benzene:methanol). Of these the procedure producing both the highest number of MS features and best signal to noise was chosen for use here. This was a modified, previously described extraction protocol that involved a solvent mixture of hexane: isopropanol (3:2) (22). To 200 µL of serum, 1.8 mL hexane: isopropanol (3:2) and 300 µL of 0.5 M KH$_2$PO$_4$ were added in a glass tube followed by vigorous vortexing. Samples were further agitated on a shaking platform for 1 h at room temperature at the speed of 80 rpm. To complete the extraction, 150 µL of water were added, mixed and centrifuged at 2000 rpm for 12 min. The upper, organic phase containing lipids was collected and dried completely under nitrogen. Dried lipid extracts were redissolved in 200 µL of chloroform: methanol (3:1) and stored at -80 °C.

**Mass spectrometric analysis of the lipid extract**

To a 20 µL aliquot of the sample extract, 23 µL of chloroform, 46 µL of methanol and 14 µL of 12 µM ammonium acetate were added. The samples were directly injected into the mass spectrometer (6230 TOF LC/MS Agilent Technologies) through an electrospray ionization (ESI) source operated in the positive ion mode. A syringe pump was utilized to inject samples at the flow rate of 2 µL/min. The ESI source employed a microspray needle having an i.d. of 50 µm. The capillary voltage was set at 3500 V. MS data was collected over mass to charge ratios (m/z) of 100-3000 with an acquisition rate of 1 spectrum/sec. Nebulizer gas and dry gas parameters were optimized to obtain a stable flow. The dry gas was set to 5 L/min at 325°C with a nebulization gas pressure of 1.03 bar. Mass Hunter-Qualitative software [Agilent Technologies] was used for data analysis. Each specimen generated a mass spectrum from which the total ion chromatogram (TIC the sum of all ion counts) from m/z 100 to 3000 was determined. A peak list having m/z values for all peaks with their abundances was generated from the mass spectrum.

To reduce analytical variability all MS peaks were normalized. For normalization, 7 abundant peaks representing different classes of lipids but showing similar abundance in both case and control sera (p>0.40) were chosen as a reference set and their abundances averaged. The m/z values and classes for
these peaks are as follows: m/z 203 (fatty acid), 369 (sterol), 666 (cholesterol ester), 758 (phosphatidylcholine (PC)), 782 (PC), 810 (PC) and 848 (triacylglycerol). These peaks were consistently seen in all the samples. The ratio of the average intensity of these 7 peaks was comparable to or more consistent than the use of TIC counts for peak normalization, but avoided the occasional high TIC for a few MS runs with a high background. Therefore, the average of the combined intensity of these 7 peaks was used to normalize all peaks across the each individual mass spectrum.

For the second confirmatory study, all the samples were processed using the same method and were directly injected onto the same MS instrument at a flow rate of 10 µL/min, using a standard sprayer having an i.d. of 120 µm. The microbore needle had previously been blocked on several occasions. The larger bore needle used here eliminated that problem. The increased flow rate and shorter run times kept the TIC, peak list and abundances comparable to the initial study. All the other parameters were kept identical. All the potential significant markers from the initial, discovery set were reanalyzed for their performance in the second confirmation set using Student’s T-test.

**Chemical characterization of the replicating lipid biomarkers**

After candidate biomarkers were found and confirmed significant by ESI-TOFMS, tandem MS (MS²) was performed to fragment and characterize all validated biomarkers. Lipid markers were fragmented using a QSTAR Pulsar 1 quadrupole (Sciex, Framingham, MA) and as needed an Agilent 6530 Q-TOF MS (Agilent, Santa Clara, CA) in the positive ion mode. For the QSTAR, samples were injected directly at a flow rate of 2 µL/min. The capillary voltage was set to 4200 V. The selection of the mass range targeted for fragmentation depended upon the m/z of the precursor ion with a spectral acquisition rate of 1 spectrum/sec. Declustering and focusing potentials were set to 65 V and 290 V respectively. Nitrogen and/or argon gas was used for collisionally-induced fragmentation. Multiple fragmentation energies were employed to obtain as complete fragmentation as possible. MS2 spectra were collected for 2 min and the
multi-channel analyzer (MCA) function was turned on resulting in summation of all 120 MS2 spectra together thus increasing signal to noise ratio.

For the Agilent 6530, sample injection was carried out at the flow rate of 10 µL/min. The capillary voltage was set to 3500 V. The drying gas flow rate and the temperature were 5 L/min and 300°C. MS2 spectra were collected from \textit{m/z} 50-3000 and the spectral acquisition rate was 3 spectra/sec. Collision energies were optimized depending upon the precursor to obtain maximum fragmentation coverage. Individual scans were summed using the add feature of the Mass Hunter program to obtain greater signal to noise. The targeted MS2 mode was used to isolate and fragment the precursor ion.

Exact mass studies were done while using a set of internal standards. Lipidmaps.org was used to tentatively determine the lipid class from the different possible classes for a particular precursor ion. Furthermore, product ions, neutral losses, fragmentation patterns of standards or fragmentation patterns found in the literature, in conjunction with predicted fragments from the databases and predicted elemental compositions based on exact mass studies were compared with the actual observed fragments to determine the class and to evaluate the possible or probable components of the individual candidate markers.

**Statistical Analyses**

For comparisons between the cases and controls, the normalized intensities of all the peaks (abundance > 200 ion counts) in the first study were subjected to a two-tailed Student’s t-test as an initial statistical comparison. A p-value <0.05 for any peak, was used to designate a potential biomarker. It is recognized that since the t-test statistic was performed for multiple peaks, separately, the true p-values are much larger than 0.05. Consequently, any candidate biomarker potentially may not be statistically significant between the cases and controls. However, using a one-at-a-time p-value calculation with a minimum threshold reduces the list of potential biomarkers. We augment this list by adding peaks that had low p-values close to 0.05 but greater than this threshold, resulting in a list of 45 potential biomarkers.
Those peaks that continued to show differential expression, as measured by a p-value of less than 0.05 as determined from a one-at-a-time t-test in the second, independent set of specimens were considered highly likely to be biomarkers. Given the goal of generating panels of biomarkers, candidates with a p-value between 0.05 and 0.10 were also included because it is recognized that some markers may be selective for a subgroup within the broader diagnosis of preeclampsia and hence may be complementary to other markers. For this reason, they were considered in potential biomarker panels and were modeled to provide better sensitivities and specificities. Biomarker panel development employed a forward selection; leave one out, logistic regression analysis approach (23-24). In modeling the MS data, there were a few peaks in a small number of specimens for which intensities could not be determined due either to ion suppression by other nearby peaks or because of a sub-threshold abundance of a specific peak in that sample. To account for these missing values, the values for these peaks were estimated using multivariate imputation by chained equations. The method of chained equations used to predict the missing values was through predictive mean matching. This uses the values of the other samples to impute missing values, allowing for all peaks to still be used (25). Using logistic regression analysis receiver operator characteristic curves were generated for these panels to allow for the determination of sensitivity (true positive) and specificity (1-true negative or false positive rate).

RESULTS

Serum lipid preeclamptic biomarkers in a discovery study

The hypothesis of this study was that there would be serum lipid biomarkers present at 12-14 wks pregnancy predictive of preeclampsia later in the same pregnancy. Using a global, serum lipidomic approach, the initial study of 27 controls and 29 cases found 45 candidate markers that were statistically or near statistically different in women who developed the disease compared with women who had uncomplicated pregnancies. The candidate markers are listed in Supplemental Table 1.
Replicating serum lipid biomarkers established in the second confirmatory study

A second confirmation study of the 45 candidate biomarkers was performed to evaluate their performance in a second set of specimens processed and analyzed comparably. This set, also collected at 12-14 wks pregnancy, involved 43 controls having uncomplicated term pregnancies and 37 cases having preeclampsia later in the same pregnancy. Of the 45 potential biomarkers, 23 continued to show a statistically significant or near significant p-values when considered one at a time. These markers are listed Table 2S.

Multi-marker panel development

Statistical modeling, using logistic regression analysis, was performed on the 23 replicating markers to develop multi-marker panels with higher predictive values. Several panels having combinations of 3-6 markers were obtained demonstrating areas under the curve (AUC) >0.85 for receiver operator characteristic (ROC) curves as illustrated in Supplemental Table 3. The ROC curves for two multi-marker combinations having AUC of 0.89 (sensitivity ~91% at a specificity of ~82% and the second ~86% and specificity ~81%) are shown in Figures 6a and 6b. A plot showing the classification of cases and controls by means of a panel of 6 markers, m/z 263.3, 383.3, 462.3, 645.5, 784.6 and 920.7 is shown in Figure 7. The overall AUC of this set is 0.88.

Chemical characterization of the validated serum lipid preeclamptic biomarkers

While absolute chemical structures of most lipids are not possible using MS, substantial chemical characterization is often provided by tandem MS fragmentation studies. Tandem MS studies were performed on all 23 confirmed serum lipid biomarkers. This provided lipid classes, exact mass studies suggested possible or probable elemental composition, fragments representing neutral losses suggested possible or probable structural components for most of the confirmed lipid biomarkers. The molecular features most consistent with the data are described in Table 2.
Fragmentation results with any chemical structural information are summarized briefly for the validated markers:

The biomarker \( m/z \) 263.2, when fragmented, demonstrated several hydrocarbon fragments in the low \( m/z \) region consistent with its having an alkane region. This marker’s fragmentation pattern was most consistent with that of a fatty alcohol or aldehyde. Its elemental composition of \( \text{C}_{18}\text{H}_{30}\text{O} \), as determined on LIPID MAPS indicates a fatty alcohol or aldehyde. This compound could also correspond to an [acyl]+ ion for linoleic acid as reported previously (26). A fragment ion at \( m/z \) 245.2, predicted to result from the (linoleoyl-H\(_2\text{O})^+ \) ion, was also observed (26) in our MS\(^2\) studies of the precursor \( m/z \) 263.2. The ions at \( m/z \) 263.2 and \( m/z \) 245.2 have also been observed in the MS\(^2\) spectrum of linoleic acid (\( m/z \) 298 for M+NH\(_4^+\)) in previous studies (26). However, the occurrence of \( m/z \) 263.2 corresponding to the (linoleoyl)+ ion was not observed in the MS\(^1\) spectrum in that report indicating that \( m/z \) 263.2 is more likely to be a protonated fatty alcohol or aldehyde and not the (linoleoyl)+ ion. The marker with \( m/z \) 301.2 was predicted to have an elemental composition of \( \text{C}_{20}\text{H}_{28}\text{O}_2 \) consistent with its being a fatty acid or fatty acid conjugate.

Another marker having \( m/z \) 383.3 produced a fragmentation pattern consistent with it being an oxidized (keto or epoxy) cholesterol (27-28). The lipid ion peak at \( m/z \) 383.3 has been identified to be dehydrated 7-ketocholesterol in a number of published studies involving human blood samples (29-31), although, we recognize that oxidation of cholesterol at some other position is possible and would result in similar fragmentation.

The species at \( m/z \) 445.4 was likely to be a cholesterol derivative as indicated by its fragmentation pattern. Fragmentation studies for two other markers having \( m/z \) 645.5 and \( m/z \) 714.6 resulted in a fragment ion at \( m/z \) 369.4, characteristic of cholesteryl esters, formed with octadecatetraenoic acid (C18:4) and docosahexaenoic acid (22:6) respectively.
To improve interpretation of fragmentation data for glycerophosphocholine (PC) lipids, studies were performed on a standard, PC (14:0/16:0). Figure 1a shows the MS² spectrum of protonated PC (14:0/16:0) having m/z 706.5 displaying a highly abundant signal at m/z 184.1 arising from the protonated phosphocholine moiety, characteristic of fragmentation of lysophosphatidylcholines (LPC), PC and SM species (32). Figure 1b displays the fragment ion at m/z 468.3 corresponding to protonated LPC (14:0/0:0) and a fragment ion at m/z 496.3 corresponding to protonated LPC (16:0/0:0) resulting from the neutral losses of the fatty acyl chains. The peak at m/z 523.4 (M+H-183)⁺ results from the loss of the entire phosphocholine head group from the precursor at m/z 706.5. These are all predicted fragments. No other fragments were seen for unpredicted truncations, combinations or any other unanticipated collision product. We also performed MS² studies on a second standard, PC (18:0/18:2) showing a protonated peak at m/z 786.6. Its fragmentation spectrum displayed fragment ions at m/z 184, 520.3 and 524.3 (Fig 2a and 2b). The signal at m/z 184 corresponds to protonated phosphocholine whereas m/z 520.3 and 524.3 correspond to LPC (18:2/0:0) and LPC (18:0/0:0) produced by neutral losses of the fatty acyl chains. Again, all fragments were exactly as predicted and no other fragments were observed. Studies involving identification of PC’s based on the neutral losses of the fatty acyl chains are commonly observed in literature (33-35). Collectively, our study and published literature suggest strongly that fragmentation patterns observed with PC lipid species follow predictable fragmentations.

Based on their fragmentation patterns, several of the lipid markers were categorized as PC lipids. The MS² fragmentation spectrum for markers with m/z values of 734.6, 760.6, 784.6, 788.6, 796.6, 798.6, 810.6, 836.6 and 895.7 displayed a prominent peak at m/z 184.07 corresponding to a phosphocholine moiety. Their elemental compositions were predicted using exact mass studies and the fatty acyl constituents were assigned based on the neutral losses from the precursor ion with help from the Lipid MS predictor feature from LIPID MAPS and following the fragment patterns observed with PC standards.
The fragmentation spectrum of the marker with $m/z$ 734.6 also displayed low abundance fragment ion with $m/z$ 496.3 suggesting the presence of 16:0 fatty acyl chain (33) (Fig 4b). Therefore, putative structure for $m/z$ 734.6 is likely to be PC (32:0) +H$^+$ having two 16:0 fatty acyl chains. The MS$^2$ spectrum of the marker with $m/z$ 760.6 displayed fragment ion peaks at $m/z$ 496.3 and $m/z$ 504.3 consistent with the presence of 16:0 and 18:1 fatty acyl chains (34). Therefore, this lipid $m/z$ 760.6 is likely to be PC (34:1) +H$^+$. Similarly, markers with $m/z$ 784.6, 788.6, 810.6 and 836.6 were predicted to be PC (36:3)+H$^+$, PC (36:1)+H$^+$, PC (38:4)+H$^+$ and PC (40:5)+H$^+$ (34). Based on the fragmentation patterns, the fatty acid components for these same markers $m/z$ 784.6, 788.6, 810.6 and 836.6 were predicted to be (16:0 and 18:1), (18:0 and 18:1), (18:0 and 20:4) and (18:0 and 22:5) respectively (34).

The MS$^2$ spectrum of marker $m/z$ 796.6 displayed a fragment ion at $m/z$ 510.3 previously reported to indicate the presence of an O-18:0 fatty acyl chain (36). Based on the several fragments observed, it is predicted to be PC(O-38:4) having a combination of O-18:0 and 20:4 fatty acyl chains consistent with the previously published literature (36).

The fragmentation spectrum of marker $m/z$ 798.6 displayed a prominent fragment ion peak at $m/z$ 780.6 likely to result from water loss from the precursor ion and is not otherwise commonly observed in the MS$^2$ spectrum of PC’s. This major neutral water loss peak observed for this precursor, suggests strongly an easily removable hydroxyl group (37). A fragment ion at $m/z$ 496.3, most likely corresponding to 16:0 fatty acid, was also observed. This fragmentation pattern most likely represents a hydroxylated PC (16:0/20:4) and this agrees with previously published literature (37). Therefore, marker with $m/z$ 798.6 was predicted to be hydroxylated PC (36:4) +H$^+$.

Due to the presence of a fragment ion at $m/z$ 184.07 and a peak indicating water loss from the precursor, the marker with $m/z$ 895.7 likely represents an oxidized sphingomyelin (SM) (odd M+H$^+$ having
even neutral mass). Oxidized SM’s have not been commonly studied using mass spectrometry (38). Therefore, the structure for m/z 895.7 could not be predicted with certainty.

Tandem MS studies were also performed on a triacylglycerol standard, TG (16:0/18:1/16:0) that displayed an ammoniated precursor peak at m/z 850.7. As shown in Figure 3, in the fragmentation spectrum of precursor m/z 850.7 (M+NH₄⁺) a fragment ion at m/z 551.5 corresponding to diacylglycerol (DG) 16:0/16:0⁺ and a fragment ion at m/z 577.5 corresponding to diacylglycerol (DG) 16:0/18:1⁺ were observed. These ions result from the neutral losses of 18:1 and 16:0 from the precursor ion. These fragments were predicted and no other fragments due to truncations, recombinations, or other transformations were observed. Based on MS² fragmentation results, the markers having m/z values of 916.8, 920.7, 928.8, 954.8, 956.8 and 958.8 represented triacylglycerols. MS² spectrum of the marker m/z 916.8 (Fig 5) displayed a fragment at m/z 881.7 likely to be due to neutral loss of water and ammonia. This suggested that m/z 916.8 is an ammoniated adduct (M+NH₄⁺) of an oxidized triacylglycerol. The peak was predicted to be epoxide of OOO_εp based on published studies of such compounds and identified by the presence of oxidized DAG fragments (39-40). For our biomarker, a fragment at m/z 617.5 corresponding possibly to epoxy diacylglycerol of OO_εp⁺ (39) and a fragment at m/z 603.5 corresponding possibly to OO⁺ (41) were observed in the MS².

Fragmentation studies of marker m/z 920.8 resulted in the production of fragment ions at m/z 575.5, 623.5 and 647.5 corresponding quite possibly to DAG fragments of 16:0/18:2⁺, 16:0/22:6⁺ and 18:2/22:6. Therefore, it is predicted to be a TG (16:0/18:2/22:6⁺)+NH₄⁺.

Fragmentation spectrum of the marker with m/z 928.8 displayed fragment ions at m/z 601.5, 613.5, 629.5 and 631.5 consistent with the fragmentation pattern of ammoniated adduct (M+ NH₄⁺) of LLL-mono-hydroperoxide (42).
MS² studies of lipids with m/z 954.8, 956.8 and m/z 958.8 resulted in water losses from their ammoniated (M+NH₄⁺) precursor ions suggesting oxidized lipids. Some other abundant fragment ion peaks represented neutral loss of fatty acids from TAG species indicating that markers with m/z 954.8, 956.8 and 958.8 belong to the class of TAG lipids. Unambiguous identification could not be made for these TAG’s because of their complicated spectra which may have contained a second overlapping isobaric or near isobaric species.

The markers with m/z 425 and m/z 462 are very likely lipids but with unknown identities due to fragmentation patterns without precedence in the literature or any of the databases.

DISCUSSION

These studies tested the hypothesis that measurement of serum lipid biomarkers early in pregnancy would identify patients at risk for developing preeclampsia later in that pregnancy. Using a ‘global’ or in-depth or ‘shotgun’ serum lipidomics approach, these studies suggest that there are predictive preeclampsia biomarkers.

Global serum lipidomic approaches are relatively new and have been applied to only a few clinical indications, e.g. Alzheimer’s disease (43-44). To our knowledge they have not been previously applied to Preeclampsia, although there have been a few reports of altered cholesterol and triglyceride profiles in preeclampsia (45-47).

Interest in lipids has increased as it has been recognized that lipids are more widely involved in cell regulatory pathways than previously thought. Lipids then may not only be altered in response to disease but it is possible that some circulating or cellular lipid species may mediate or contribute to aspects of disease. Given the uncertainties in the etiology and prediction of preeclampsia, there is interest in both
developing better assessments of preeclampsia risk as well as better understanding the early changes that precede the fully manifest disorder.

Our approach led to the detection and confirmation of 23 unique lipid biomarkers for preeclampsia. It is entirely possible that these circulating factors reflect the consequence of early disease, but some may have a more direct biological role. In the absence of an accepted animal model for preeclampsia, early events leading to this disease are poorly understood. A number of lipidomic biomarkers based on their lipid class and other potential characterization could be related to pathological processes linked to preeclampsia. Thus, considering the changes in these specific biomarkers may be useful in providing biochemical information about early antecedents of this disease. Based on previous literature, some are at least suggestive. For example, among the several markers, the peak at $m/z$ 383.3, was predicted to be an oxidized cholesterol (keto or epoxy). 7-ketocholesterol, which has the same mass, has been proposed to contribute to atherosclerosis (48) and vascular changes similar to atherosclerotic disease have been reported in preeclampsia (49). This might explain the higher levels of this marker observed in women with later preeclampsia.

The candidate biomarkers having $m/z$ values 734, 760, 784, 796, 810 and 836 were found to be higher in the serum of preeclamptic women and belong to the lipid class of glycerophosphocholines (PC). Placental ischemia and apoptosis have been reported in preeclampsia, resulting in cell lysis with release of membrane constituents, including likely phosphatidylcholines, into the circulation (50-51). This might explain the higher levels of these markers in preeclamptic cases.

The marker with $m/z$ 798 was predicted to be an oxidized (hydroxylated) PC (16:0/20:4) as indicated by the fragmentation pattern seen and consistent with other studies (37). The exposure of PC species to reactive oxygen species (ROS) can result in oxidation of the species. There have been several
reports of increased production of ROS in preeclampsia, which might explain the increased production of this marker in preeclamptic cases (52).

Currently, there are no accepted biomarkers for predicting preeclampsia. A number of candidates have been proposed, especially pro-angiogenic and anti-angiogenic factors, and while they have repeatedly demonstrated changes in many women with established disease, they have shown poor sensitivity and specificity in predicting all forms of preeclampsia (11-13). Therefore, there continues to be a need for predictive biomarkers of this disorder with more reliable sensitivities and specificities. The AUC values of 0.89 for two panels of markers detected in our study are substantially better than the previously reported values in the literature. The utility of these markers across two studies suggests these they should be considered of interest, especially given their being found early in the pregnancy and predicting events months later.

Utilization of this direct lipidomic approach as used here could provide a high throughput method for analysis of individual lipid species from diverse classes without a chromatographic separation step. This method provides comparative quantitation of species and with standards could allow for absolute quantitation. MS methods are currently in use in the clinical laboratory, e.g. drugs of abuse analysis. In addition, our approach allows for chemical characterization of interesting lipids as well as analysis analogous to multiple reaction monitoring as used for peptides and proteins. This method can have a linear range above 1000, even over the low concentration range, making it efficient for studying low abundant lipid species (53).

The approach analyzed ~2,000 lipids but could also be carried out in the negative ion mode providing an even more comprehensive evaluation of serum lipids were it needed. This expansion of scope would likely provide additional useful lipid biomarkers for early diagnosis of preeclampsia. The current
approach represents a versatile method requiring relatively simple specimen preparation and demonstrating adequate reproducibility. As such, it could be applied to the analysis of blood samples for potentially many other clinical targets.

While the current studies are suggestive, the specimens available were primarily from Caucasian subjects. It is recognized that additional studies are needed to determine whether the markers found are more universal in their application. Certainly, more serum lipidomic studies are needed to demonstrate clinical utility. Nevertheless, global serum lipidomics is a valuable and robust approach for the discovery and identification of lipid biomarkers for disease and may have provided useful predictive markers for preeclampsia.

Acknowledgements

We would like to thank members of the staff from the Department of Obstetrics and Gynecology, University of Utah School of Medicine and from Intermountain Health Care and its associated clinics for their help with specimens. We would like to thank the Deseret Foundation for providing a portion of the funding used for this work. Additionally, we received a CHIRP grant from the College of Physical and Mathematical Sciences, Brigham Young University supporting this research.
REFERENCES


Figure Legends

Figure 1. MS\(^2\) spectrum for a glycerophosphocholine standard PC (14:0/16:0) with \(m/z\) 706.5 (M+H\(^+\)).

a) The abundant fragment ion at \(m/z\) 184 corresponds to the protonated phosphocholine head group.

b) The fragment at \(m/z\) 468.3 corresponds to protonated LPC (14:0/0:0). Similarly the fragment at \(m/z\) 496.3 corresponds to protonated LPC (16:0/0:0). Neutral loss of the entire phosphocholine head group leads to the formation of the ion at (M+H-183)\(^+\) at \(m/z\) 523.

Figure 2. MS\(^2\) spectrum for a glycerophosphocholine standard PC (14:0/16:0) with \(m/z\) 786.6 (M+H\(^+\)).

a) The abundant fragment ion at \(m/z\) 184 corresponds to protonated phosphocholine head group. Neutral loss of the entire phosphocholine head group leads to the formation of the ion at (M+H-183)\(^+\) at \(m/z\) 603.5.

b) The fragment at \(m/z\) 520.3 corresponds to protonated LPC (18:2/0:0). Similarly the fragment at \(m/z\) 524.3 corresponds to protonated LPC (18:0/0:0).

Figure 3. MS\(^2\) spectrum for a triacylglycerol standard TG (16:0/18:1/16:0) having \(m/z\) 850.7 (M+NH\(_4^+\)).

The fragment ion at \(m/z\) 551.5 corresponds to diacylglycerol 16:0/16:0\(^+\) resulting from the neutral loss of 18:1 from the precursor. Similarly, the fragment ion at \(m/z\) 577.5 corresponds to diacylglycerol 16:0/18:1\(^+\) resulting from the neutral loss of 16:0 from the precursor ion.

Figure 4. a) MS\(^2\) spectrum for a lipid marker with \(m/z\) 734.5. Fragmentation suggests that it is likely to be a protonated adduct of PC -32:0. b) The fragment ion at \(m/z\) 496.3 is likely due to protonated LPC (16:0/0:0) resulting from the neutral loss of another 16:0 from the precursor ion.

Figure 5. MS\(^2\) spectrum for a lipid marker with \(m/z\) 916.8. It was predicted to be an ammoniated adduct of OO\(_{ep}\). The fragment ions at \(m/z\) 617.5 and \(m/z\) 603.5 are likely to be OO\(_{ep}\)^+ and OO\(^+\).
Figure 6. a) Resulting receiver operator characteristic curve generated by logistic regression analysis that included 6 lipid biomarkers having mass to charge ratios of 383, 445, 784, 796, 798 and 920. The AUC was 0.89 with a sensitivity of ~91% at a specificity of ~82%. b) Resulting receiver operator characteristic curve generated by logistic regression analysis of another combination of 6 lipid biomarkers having mass to charge ratios of 263, 383, 445, 645, 784 and 916. Some peaks are common to the two panels. The AUC of this set was also 0.89 with a sensitivity of ~86% at a specificity of ~81%.

Figure 7. A plot showing the ability of a panel of 6 biomarkers, m/z 263.3, 383.3, 462.3, 645.5, 784.6, 920.7, to correctly classify cases and controls. Controls are shown in blue and cases are displayed in red. Each point represents a different patient. The overall AUC is 0.88 for this set of markers.
Table 1a: Demographics for the First, Discovery Study

<table>
<thead>
<tr>
<th>Maternal age (yrs)</th>
<th>Gestational age (wks)</th>
<th>BMI (kg/m²)</th>
<th>SBP</th>
<th>DBP</th>
<th>Birthweight (gm)</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>30 ± 4.6</td>
<td>36.4 ± 3.07</td>
<td>30.8 ± 5.42</td>
<td>149.5 ± 17.4</td>
<td>90.1 ± 12.3</td>
<td>2755 ± 801</td>
</tr>
<tr>
<td>Controls</td>
<td>27 ± 4.8</td>
<td>39 ± 1.23</td>
<td>30.4 ± 5.6</td>
<td>138.6 ± 14.7</td>
<td>80 ± 12.5</td>
<td>3459 ± 349</td>
</tr>
</tbody>
</table>

Note: BMI = body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; Cau = Caucasian

Table 1b: Demographics for the Second, Confirmatory Study

<table>
<thead>
<tr>
<th>Maternal age (yrs)</th>
<th>Gestational age (wks)</th>
<th>BMI (kg/m²)</th>
<th>SBP</th>
<th>DBP</th>
<th>Birthweight (gm)</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>29 ± 4.7</td>
<td>36.3 ± 3.2</td>
<td>31.7 ± 5.89</td>
<td>147 ± 20.8</td>
<td>90 ± 15.1</td>
<td>2817 ± 860.9</td>
</tr>
<tr>
<td>Controls</td>
<td>26 ± 4.7</td>
<td>38.7 ± 1.6</td>
<td>27.9 ± 4.7</td>
<td>117 ± 14.4</td>
<td>70.4 ± 8.2</td>
<td>3295 ± 535.2</td>
</tr>
</tbody>
</table>

Note: BMI = body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; Cau = Caucasian
Table 2. Chemical Characterization of Preeclamptic lipid Biomarkers

<table>
<thead>
<tr>
<th>No</th>
<th>m/z</th>
<th>Adduct</th>
<th>Class</th>
<th>Possible identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>263.2</td>
<td>M+H*</td>
<td>Fatty alcohol and aldehydes</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>301.2</td>
<td>M+H*</td>
<td>Fatty acids and conjugates</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>383.3</td>
<td>M*</td>
<td>Oxidized cholesterol</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>425.1</td>
<td>------</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>445.4</td>
<td>M+H*</td>
<td>Cholesterol derivatives</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>462.3</td>
<td>------</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>645.5</td>
<td>M+H*</td>
<td>Cholesteryl esters</td>
<td>C18:4 cholesteryl ester</td>
</tr>
<tr>
<td>8.</td>
<td>714.6</td>
<td>M+NH₄</td>
<td>Cholesteryl esters</td>
<td>C22:6 cholesteryl ester</td>
</tr>
<tr>
<td>9.</td>
<td>734.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-32:0 16:0 and 16:0</td>
</tr>
<tr>
<td>10.</td>
<td>760.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-34:1 16:0 and 18:1</td>
</tr>
<tr>
<td>11.</td>
<td>784.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-36:3 18:1 and 18:2</td>
</tr>
<tr>
<td>12.</td>
<td>788.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-36:1 18:0 and 18:1</td>
</tr>
<tr>
<td>13.</td>
<td>796.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-O-38:4 O-18:0 and 20:4</td>
</tr>
<tr>
<td>14.</td>
<td>798.6</td>
<td>M+H*</td>
<td>Oxidized Glycerophosphocholine</td>
<td>PC-36:4+OH 16:0 and 20:4</td>
</tr>
<tr>
<td>15.</td>
<td>810.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-38:4 18:0 and 20:4</td>
</tr>
<tr>
<td>16.</td>
<td>836.6</td>
<td>M+H*</td>
<td>Glycerophosphocholine</td>
<td>PC-40:5 18:0 and 22:5</td>
</tr>
<tr>
<td>17.</td>
<td>895.7</td>
<td>M+H*</td>
<td>Oxidized sphingomyelin</td>
<td></td>
</tr>
</tbody>
</table>
18. 916.8 M+NH₄ Oxidized triacylglycerol OOOepoxide

19. 920.7 M+NH₄ Triacylglycerol TG-56:8 16:0, 18:2 and 22:6

20. 928.8 M+NH₄ Triacylglycerol LLL-mono-

21. 954.8 M+NH₄ Triacylglycerol hydroperoxide

22. 956.8 M+NH₄ Triacylglycerol

23. 958.8 M+NH₄ Triacylglycerol
Figure 3
Fig 6a and 6b
Figure 7