HHS Public Access

Author manuscript

ACS Appl Mater Interfaces. Author manuscript; available in PMC 2024 August 16.

Published in final edited form as:

ACS Appl Mater Interfaces. 2023 August 16; 15(32): 38185–38200. doi:10.1021/acsami.3c04260.

First trimester prediction of preterm birth in patient plasma with machine-learning guided Raman spectroscopy, and metabolomics

Lilly Synan[a,b,†], **Saman Ghazvini**[a,b,†], **Saji Uthaman**[a,b], **Gabriel Cutshaw**[a,b], **Che-Yu Lee**[c], **Joshua Waite**[d], **Xiaona Wen**[b], **Soumik Sarkar**[d], **Eugene Lin**[c], **Mark Santillan**[e], **Donna Santillan**[e], **Rizia Bardhan**[a,b,*]

[a]Department of Chemical and Biological Engineering, Iowa State University, Ames, IA 50012, USA

[b]Nanovaccine Institute, Iowa State University, Ames, IA 50012, USA

[c]Department of Chemistry and Biochemistry, National Chung Cheng University, Chiayi 62106, Taiwan

[d]Department of Mechanical Engineering, Iowa state University, Ames, IA 50012, USA

[e]Department of Obstetrics and Gynecology, Carver College of Medicine, University of Iowa, Hospitals & Clinics, Iowa City, IA 52242, USA

Abstract

Preterm birth (PTB) is the leading cause of infant deaths globally. Current clinical measures often fail to identify women who may deliver preterm. Therefore, accurate screening tools are imperative for early prediction of PTB. Here we show Raman spectroscopy is a promising tool for studying biological interfaces, and we examine differences in the maternal metabolome of the first trimester plasma of PTB patients and those that delivered at term (healthy). We identified fifteen statistically significant metabolites that are predictive of onset of PTB. Mass spectrometry metabolomics validates the Raman findings identifying key metabolic pathways that are enriched in PTB. We also show that patient clinical information alone and protein quantification of standard

*Corresponding author: rbardhan@iastate.edu.
†Equal contributing first authors

Author contributions

RB conceived the project idea and directed the study design. LS and SG conducted all Raman spectroscopy and MS experiments, analyzed them, and wrote the manuscript together. SU measured all the ELISA in patient samples and edited the manuscript. GC assisted with AUC-ROC analysis of Raman data and Raman + clinical data, helped with revision experiments, and edited the manuscript. SS and JW provided custom MATLAB codes for running tSNE and data separation. CYL and ECL provided the MATLAB codes for Raman data processing. XW trained in experimental techniques, provided discussions on Raman data, and edited the manuscript. MS and DS provided patient samples, clinical data, discussions of data relevance to pregnancy, and edited the manuscript.

Conflicts of interest

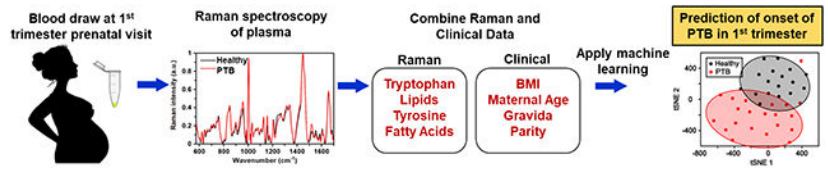
The authors declare no conflict of interest.

Supporting Information

Clinical information tables for individual healthy and PTL patients; Pearson's correlation heatmap values; Raman metabolites associated with each KEGG metabolic pathway; Raman spectra of individual healthy and PTL patients; Cross validation of SVM model; Box plots of additional Raman metabolites; and LC-MS metabolite heatmap. This material is available free of charge via the Internet at http://pubs.acs.org.

inflammatory cytokines, both fail to identify PTB patients. We show *for the first time* that synergistic integration of Raman and clinical data guided with machine learning results in an unprecedented 85.1% accuracy of risk stratification of PTB in first trimester that is currently not possible clinically. Correlations between metabolites and clinical features highlight body mass index and maternal age as contributors of metabolic rewiring. Our findings show that Raman spectral screening may compliment current prenatal care for early prediction of PTB, and our approach can be translated to other patient-specific biological interfaces.

Graphical Abstract



Keywords

Raman spectroscopy; preterm birth; pregnancy; metabolomics; machine-learning; metabolism

1. Introduction

Spontaneous preterm birth (PTB) occurring at <37 weeks of gestation is the leading cause of neonatal mortality worldwide, and results in >1 million childhood mortalities under the age of 5.[1] The mechanisms that lead to PTB are poorly understood due to the heterogeneities associated with the many risk factors such as maternal age, race, etc. Whereas many risk-scoring systems based on a patient's epidemiologic and obstetric factors have been applied to predict PTB, sensitivity and positive predictive values have been low and failed to identify women at risk of PTB; patients identified as high risk often did not deliver preterm.[2] Current clinical standards that identify women at high risk of PTB include cervical length assessment and screening for fetal fibronectin, a protein elevated in the vaginal pool near birth.[3] However, these approaches are inadequate and often inconclusive in predicting the complex phenotype of PTB.[4, 5] Despite advances in biomarker discovery, prediction of the onset of PTB has remained a challenge because (1) current clinical assays have poor sensitivity and specificity which fail to identify relevant biomarkers;[6] (2) single biomarker tests are ineffective due to low positive predictive values;[7] (3) there are significant inter- and intra-assay disparities due to variations in detection method, sample type used, and timing of sample collection which make comparative analysis of biomarkers difficult for PTB prediction;[8] and (4) while some studies have reported multiplexed screening using the same sample type and platform,[9] there were significant variations in the study designs and clear conclusions could not be drawn. An unmet clinical need exists for an effective screening tool that enables quantitative detection of unconventional biomarkers of PTB with high sensitivity and specificity at clinically relevant levels, yet simultaneously be low-cost allowing early, accurate, and affordable screening for patients.

Recent findings support dynamic metabolic changes in the placenta that are critical for uterine function, nutrient transport to the fetus, and fetal growth and development.[10] Therefore, deficits in the capacity of the placenta to maintain this dynamic and precisely programmed metabolic stability during pregnancy could contribute to PTB. Mass spectrometry (MS) remains a gold standard in metabolomics, offering high specificity, selectivity, and multiplexed measures of metabolites in samples. However, MS is also cost-prohibitive and time-intensive, requiring an elaborate extraction process that is prone to user error. MS also requires larger sample volumes (~60 μL of serum/plasma), and it is destructive, i.e., the sample is consumed during the measurement. Raman spectroscopy is an optical technique where incident photons induce vibrations in the samples, manifesting as inelastic light scattering, that reveals biochemical information as a unique "fingerprint" of that sample.[11] Raman is a low-cost, highly sensitive approach that enables rapid analysis (~30 min per sample) in very small sample volumes (~3 μL) capturing the chemical footprint of multiple metabolites in a high-throughput format. Raman is non-destructive, i.e., a sample can be archived and measured multiple times. It offers an extraction-free approach where serum, plasma, or other biofluids can be directly measured without further preparation. Our group has leveraged the strengths of Raman both as a label-free approach to assess cellular metabolism,[12] as well as a labeled surface-enhanced Raman approach utilizing nanoparticles to measure biomarkers in *in vivo*,[13, 14] and in sera.[15]

Here, we demonstrate the utility of Raman in metabolic measurement of plasma from first trimester pregnant patients who delivered at term (healthy) and those who delivered preterm (<37 weeks). We integrated machine learning models with Raman data to distinguish healthy patients and those at high risk of PTB and identified key metabolites correlated to the gestational age at delivery. We also leverage the merits of MS metabolomics to validate our findings in Raman and demonstrate synergism between the two approaches in metabolites measured and the corresponding metabolic pathways enriched using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. We demonstrate that analysis of standard proinflammatory cytokines, which have been implicated in the pathogenesis of PTB, with enzyme-linked immunosorbent assay (ELISA) is neither conclusive nor predictive of the onset of PTB in the first trimester consistent with literature findings. Finally, based on measurements acquired from the first trimester, we report that Raman spectral analysis when synergistically combined with patient clinical information and obstetric history, offers an accurate prediction of patients who will subsequently have PTB (Scheme 1). Our results indicate that Raman-based metabolic profiling has the potential to enable a robust and clinically relevant screening tool that may predict those at risk of PTB and allow for the appropriate risk stratification and surveillance.

2. Results and Discussions

This study evaluated total 37 plasma samples from the first trimester of pregnant patients who delivered at term (healthy) and those who had spontaneous preterm birth (PTB). The maternal demographic data is shown for healthy and PTB patients (Table 1). Table 1 shows the maternal age of patients, body mass index (BMI), gestational age when they delivered, number of pregnancies (gravida), number of births (parity), prior pregnancy loss and known comorbidities. Diabetes was a common comorbidity among the 37 participants. Average

prior spontaneous pregnancy losses were also higher in PTB patients, a known risk of preterm birth. The clinical information for individual healthy and PTB patients are presented in Tables S1 and S2 in the supporting information. The average gestational age when plasma was collected for the study was 12.1 ± 1.2 weeks with a high of 15.7, low of 10.1, and median of 12 weeks.

We measured the Raman spectra of the 37 plasma samples using a 120 mW, 785 nm laser with a spot size of 1 μm resolution and a 50 × objective. The plasma samples (3 μL) were deposited on CaF_2 disks, dried for 20 min, and immediately measured. The Raman data was normalized with the standard normal variate (SNV) method in a custom Matlab code, and once processed, all data was renormalized so that the 1447 cm^{-1} lipid peak would have a consistent value of 1. We chose this peak for normalization as it had minimal variation across samples. The normalized spectra of healthy and PTB patients (Figure 1a, Figure S1) shows significant differences in their molecular makeup, which are highlighted in the difference spectrum (Figure 1b). The difference spectrum was obtained by subtracting the healthy Raman spectrum from the PTB spectrum. Positive values in Figure 1b indicate an increase in a specific Raman peak in PTB patients, which suggests an increase in an abundance of a specific metabolite. Likewise, negative values indicate metabolites that are lower in PTB patients. The Raman peaks and the tentative metabolic assignment are listed in Table 2. In assigning our Raman peaks, we used a three step approach: (1) we referred to the literature we have cited in Table 2; these are a collection of highly cited and well-established Raman reference papers[16–20]; (2) we referred to published literature focused on preterm labor (see ref. 8, 26–46) to ensure that the metabolites assigned are relevant in pregnancy; and (3) we also referred to published literature that indicate metabolites present in blood and blood components.[21] This three-step approach gave us confidence in our tentative peak assignments, and the metabolic pathways that have resulted from these assignments. For metabolites with multiple peaks (such as some of the amino acids in Table 2), we only considered the strongest Raman peak of that metabolite as smaller peaks may not appear in a complex media such as plasma.

The peaks were analyzed with student *t*-tests to evaluate statistical significance between healthy and PTB patients which indicated 15 significant metabolites differentiate the two patient cohorts. We employed t-distributed neighbor embedding (tSNE), an unsupervised machine learning based on nonlinear dimension reduction approach. Our tSNE analysis of Raman data (Figure 1c) shows separation between healthy and PTB patient samples in the first trimester. Here, each point represents a single patient's averaged and normalized spectrum. The axes are unitless, representing a nonlinear combination of features in the data. An exaggeration value of 2 with 30 PCA components and a perplexity value of 15 were used to generate the tSNE plot.[22] The accuracy of the Raman data and the separation observed in tSNE was confirmed with an area under the curve - receiver operating characteristic curve (AUC-ROC) analysis (Figure 1d). The ROC curve determines the probability of correctly classifying data based on the sensitivity and specificity of the corresponding Raman peaks. The sensitivity is the percentage of patients with PTB (true positive rate), and the specificity is the percentage of healthy patients (true negative rate). A weighted support vector machine (SVM) algorithm was used to generate the AUC-ROC curves using MetaboAnalyst's Biomarker Analysis module. SVM is a versatile approach for the binary classification of

samples and regression analysis, and it is often combined with Raman data for biological classification.[23–25] The data scored an AUC of 0.839 with a 95% confidence interval (CI) of 0.6–1 using statistically significant peaks of Raman. For each patient sample, we used a mean spectrum that was averaged from 100 RS measurements per sample. When building the SVM model, Metaboanalyst uses a 2/3 cross-validation split. The cross-validation is then conducted 100 times on the dataset, giving a smooth AUC-ROC and a 95% confidence interval from the cross-validation process. The cross-validation accuracy characterization is provided in the supporting information (Figure S2) with the list of the peaks used to build the SVM model. Our AUC results demonstrate that Raman spectroscopy has the potential to classify patients accurately during the first trimester who are likely to have PTB, which is currently not possible clinically.

We show twelve statistically significant Raman peaks in Figure 2, and additional peak analysis is provided in the Supporting Information (Figures S3). Metabolites identified in Raman are directly correlated to the pathophysiology of PTB. We observed an increase in DNA (668 cm^{-1}) that may be correlated to DNA methylation, an epigenetic mechanism associated with the risk of PTB.[26, 27] Alterations in the intrauterine environment changes DNA methylation patterns at the cytosine-guanine dinucleotide sites promoted by multiple factors, such as diet, smoking, and inflammation. Other studies have also shown that Raman peaks of DNA increase with methylation.[28] An increase in DNA may also be correlated to cell-free fetal DNA in the maternal circulation that typically increases at the onset of birth at term and has also been associated with PTB.[29, 30] We also observe an increase in histidine (659 cm^{-1}) and methionine (681 cm^{-1}), both of which are essential amino acids (AAs) that play a critical physiological role in scavenging reactive oxygen and nitrogen species. Yet an increase in these AAs is correlated to oxidative stress and inflammation,[31] which are two detrimental factors in PTB pathology.[8] Increase in the methionine pathway has been correlated to PTB in other studies as well.[32] We observe that metabolites of the tricarboxylic acid (TCA) cycle citric/succinic acid (940 cm^{-1}) increase in our PTB cohort. TCA cycle metabolites have been associated with oxidative stress and are well documented in PTB.[33] Next we find that peaks associated with carotenoids (1154 and 1517 cm^{-1}) is also elevated. β-carotene is naturally found in fruits and vegetables and not endogenous to humans; β-carotene serves as a precursor to vitamin A biosynthesis.[34] Vitamin A is important for fetal health, and deficiencies are associated with PTB.[35] Here, an increase in β-carotene may be associated with the diet of the PTB patients (or intake of supplements). We note that in our recent study focused on Raman spectral analysis of preeclampsia patients throughout pregnancy, we also observed an increase in carotenoids in the preeclamptic cohort in all three trimesters.[36] This suggests that in addition to patient diet, a correlation of β-carotene concentrations in inflammatory phenotypes. However, we aim to investigate this further, in a future study with a larger cohort of patient samples. We find peaks corresponding to lipids and fatty acids (FAs) including myristic acid (956 cm^{-1}), long chain FAs (1173 cm^{-1}), and phospholipids including phosphatidylcholine/phosphatidylethanolamine (PC/PE, 1656 cm^{-1}) also increase in PTB. An increase in this group of metabolites in maternal plasma has been associated with premature rupture of membranes and increased risk of preterm labor.[37] Indeed, complex crosstalk is at play between maternal lipids and signaling molecules secreted from the placenta,[38] and an increase in membrane phospholipids, including PC/PE,

may be a consequence of high inflammation in PTB patients.[39] Sphingolipids have also been associated with increased prostaglandin levels which influence dilation and contractions during labor.[40]

Among the metabolites that decrease in PTB patients include AAs such as tryptophan (1551 cm^{-1}), tyrosine (1615 cm^{-1}), and phenylalanine (1583 cm^{-1}, Figure S1), which have been directly linked to parturition and PTB. These AAs are essential building blocks in protein synthesis, as they are actively transported across the placenta from mother to fetus. As a patient nears term, a continued increase in fetal AA metabolism depletes maternal AAs.[41] From our findings, a decrease in AAs in the first trimester of PTB patients may be a metabolic biomarker of dysregulated maternal-fetal tolerance and premature placental senescence leading to PTB. Among AAs, the phenylalanine-tyrosine metabolic relationship drives many processes key to healthy fetal development. Phenylalanine, an essential AA, is obtained through diet and intracellularly converted to tyrosine, a non-essential amino acid. Tryptophan, also an essential AA, plays a crucial role in many metabolic functions necessary for fetal growth, placental function, and immune regulation.[42] This suggests that a decrease in these amino acids in the first trimester is an early indicator of cervical changes and placental inflammation leading to PTB.[43,] A recent multi-omics study confirmed the role of these key AAs and corresponding pathways in the inflammatory pathophysiology of PTB.[44] We also observed a statistically significant decrease in the Amide III peak (1249 cm^{-1}) representing proteins, including collagen-like and other proteins.[45] A decrease in Amide III among the PTB patient cohort directly correlates to reduced AAs, which intricately participate in protein synthesis. Proteins are a key component of the extracellular matrix and are critical to the maternal-fetal interface providing structural integrity to the uterine endometrium and decidua in pregnancy.[46] Deficiencies in proteins in the first trimester could indicate premature cervical ripening and parturition, increasing the risk of PTB. We used correlation analysis to detect patterns in statistically significant metabolites with respect to the gestational age at delivery in patients (Figure 2m). For this analysis, the Raman peak values of all patient samples (n=17 healthy and n=20 PTB) and their corresponding gestational age was uploaded to the Statistical Analysis [one factor] module of MetaboAnalyst. The module used the Pearson's correlation test to determine positive and negative values that associate peaks to gestational age.

A Pearson product-moment coefficient (r) was specified for the correlation measure. Negative values indicate that a metabolite decreases in abundance in PTB, while positive values designate an increase in abundance in PTB. Positive and negative values also imply that gestational age increases or decreases jointly with changes in healthy patients or inversely in PTB patients. From top to bottom, metabolites increase in correlation strength. Values between ± (0 – 0.29) suggest a weak correlation, ± (0.30 – 0.50) moderate correlation, and ± (0.51 – 1) suggest strong correlation of variates to each other.[47, 48] Our results show methionine (r = −0.41187), DNA (r = −0.39267), long-chain FAs (r = −0.36687), phenylalanine (r = 0.36665), NADH (r = 0.32524), tryptophan (r = 0.30092) and lipids/phospholipids (r = −0.30) have a moderate correlation to gestational age at delivery when measured in the first trimester. The remaining metabolites have weak correlation values. The nicotinamide adenine dinucleotide (NAD+) and its reduced form NADH are redox couples that are essential in energy metabolism including glycolysis and

mitochondrial oxidative phosphorylation. NAD+/NADH regulate the cellular redox state, and the loss of redox homeostasis of these molecules has been linked to loss of pregnancy.[49] These data suggest that targeted screening of these metabolites in the first trimester may be predictive for patients at risk of PTB.

We validated our Raman findings with mass spectrometry (MS) metabolomics using GC-MS and LC-MS/MS on plasma extracts to confirm the differences between PTB and healthy patients in the first trimester (Figure S4). MS is a gold standard in metabolic profiling, identifying specific metabolites with high selectivity. However, larger sample volumes are required for MS of plasma (60 μL vs. only 3 μL for Raman) and a time- and labor-intensive sample extraction process is needed (no sample prep necessary for Raman). The detailed procedure for metabolite extraction from plasma for MS, and sample analysis is provided in the methods. The metabolites assessed with MS are represented via a heat map (Figure 3a) and quantitative analysis (Figure S5) and show an increase in a number of membrane phospholipids including PC/PE as well as in lipids and fatty acids for PTB patients reflecting the trends observed in Raman metabolic analysis (Figure 2). Metabolites from the TCA cycle such as citric acid and other metabolites in MS also show an increase in PTB which aligns well with an increase in Raman peaks of citric/succinic acid (940 cm^{-1}, Figure 2d). In MS, some sugars show increase while others show decrease indicating that the overall metabolite class of sugars may not have a significant change between PTB and healthy cohort early in pregnancy; similar trends were observed in Raman peaks for sugars (Figure S2). Among the class of AAs, we find a decrease in β-alanine in PTB cohort, minimal change in glutamic acid between the two cohorts, and an increase in homomethionine. These trends are analogous to changes in amino acids observed in Raman and also highlights the complex interplay of AAs in protein synthesis necessary for fetal development vs. their role in oxidative stress which may lead to adverse effects in pregnancy. Note that some variations between trends observed in Raman and MS is expected as MS identifies a much larger subset of AAs than Raman. These variations may be correlated to how samples are measured in the two techniques. Raman measures the inelastic scattering of photons induced by the vibrational modes in samples where molecules with stronger vibrational motions (such as fatty acids) have intense peaks in the spectra. MS utilizes a mass-to-charge ratio to identify chemical and structural information, enabling a higher specificity in distinguishing metabolites, although overlap with irrelevant molecules often occurs depending on the databases used for metabolite identification. To identify metabolites that overlap between MS and RS and those that distinctly appear in one or the other approach, we present Venn diagrams (Figure 3b) that demonstrate that RS and MS combined together provide a more comprehensive metabolic information. Therefore, we view that MS and Raman are not competitive rather highly complementary where the strengths of each approach is combined together to elucidate metabolic rewiring during the pathophysiology of PTB.

The metabolites tentatively identified with Raman and validated with MS were then used to examine possible metabolic pathways that may be responsible towards the onset of PTB (Figure 4a, b). We used the KEGG pathway analysis obtained with MetaboAnalyst's Enrichment Analysis module (see methods for details on how MetaboAnalyst was used). Table S3 shows the KEGG pathways that are associated with each metabolite found in

Raman. Group labels such as DNA, lipids, fatty acids, and Amide I and III are generic and cannot be used for enrichment analysis. The enrichment analysis provides 18 enriched pathways in Raman that overlap with MS. The metabolites that correspond to the Raman peaks (Figure 2, S3, and Table S3) and are found in the MS heat map (Figure 3, S4) are directly involved in the pathways listed. We also found several pathways that are observable in Raman but are indirectly correlated to pathways in MS. For example, the pantothenate and CoA biosynthesis pathway in MS where pantothenate (also known as vitamin B5) is absorbed in the intestine from food, and gut microbiota serves as the main precursor to the biosynthesis of coenzyme A (CoA).[50] This pathway is involved in the synthesis and degradation of fatty acids, including phospholipids, gangliosides, and bile acids among others, and it is directly associated with Raman metabolic pathways of fatty acid biosynthesis, degradation, elongation, and bile acid biosynthesis. The nicotinate and nicotinamide metabolism in MS is another example where nicotinamide, also known as niacin or vitamin B3, is biosynthesized from tryptophan and participates in the nicotinamide adenine dinucleotide (NAD+/NADH) redox cycle.[51] The tryptophan-niacin pathway is associated with tryptophan metabolism in Raman enrichment analysis, and as discussed previously (Figure 2), tryptophan correlates to the inflammatory pathophysiology of PTB. Finally, aminoacyl-tRNA biosynthesis and pyrimidine metabolism observed in both MS and Raman metabolic pathways are linked with purine metabolism seen in MS enrichment analysis, and correlates to changes in the DNA peak in Raman (Figure 2a). Indeed, DNA methylation associated with PTB, as discussed previously in Figure 2, can be measured by the stoichiometric ratio of purine and pyrimidine bases.[52] Tyrosine, tryptophan, and phenylalanine are also essential to the aminoacyl-tRNA metabolism, and collectively, these metabolites support fetal health and development and correlate to maternal parturition. The pentose phosphate pathway (PPP) in MS is also associated with pathways found in Raman. PPP proceeds in parallel to glycolysis and they share the starting molecule, glucose-6-phosphate. PPP produces NADPH, which is then utilized in the metabolic pathways associated with biosynthesis of steroids and fatty acids, both of which are present in Raman KEGG analysis. A more expansive enrichment pathway in MS is expected since MS measures a larger subset of metabolites. Our results demonstrate that the tentative metabolites identified with Raman and the possible pathways determined from KEGG may be potential early predictors of PTB.

Next, we determined if synergistic integration of Raman metabolic profiling with patient clinical history, which is a routine part of prenatal care, would enable a more accurate early diagnosis of PTB than achieved with clinical information alone or Raman data alone. We normalized patient demographic data, (Table 1, Tables S1, S2) including maternal age, BMI, gravida, parity, and previous pregnancy losses for tSNE and the correlation heat map (Figure 5). The data were normalized by dividing by the maximum value of a particular clinical parameter within both PTB and healthy groups. For example, the BMI of all patients was divided by the highest BMI among both patient cohorts to get a maximum of 1 and all other values being a fraction. This normalization was performed to aid in tSNE data visualization of clinical parameters. The tSNE in Figure 5a was generated from the normalized patient data; tSNE parameters included an exaggeration value of 2, 2 PCA components, and a perplexity of 14. This plot demonstrates that clinical information alone cannot accurately

classify patients at high risk of PTB in the first trimester. Gestational age was excluded from the tSNE of clinical data as that information would not be available in the first trimester. The combination of Raman metabolic classification (Figure 5b, also shown in Figure 1c) with the normalized clinical data of patients (Figure 5c) enables a robust separation of the two patient cohorts better than can be achieved with either dataset alone. We generated AUC-ROC curves of the Raman + clinical data shown in Figure 5c using an SVM algorithm; the AUC was based on statistically significant Raman peaks and the five clinical parameters included in Figure 5a. We achieved an AUC of 0.851 with a 95% CI of 0.638–1 (Figure 5d) highlighting that Raman metabolic analysis combined with patient clinical history enables an unprecedented diagnostic accuracy in the 1st trimester to identify patients at high risk of PTB. Please see Figure S2b for the cross validation of the SVM model that enabled the AUC-ROC analysis.

We examined if metabolites measured in the first trimester correlated with patients' clinical information (not including gestational age at delivery). We generated a correlation heat map where blue cells represent an inverse correlation and red cells represent joint correlation (Figure 5e). Pearson's product-moment coefficient (r) values for each pairing are provided in the Supporting Information (Table S4) where values between ± (0 – 0.29) suggest a weak correlation, ± (0.30 – 0.50) moderate correlation, and ± (0.51 – 1) suggest strong correlation of metabolites and clinical parameters. Our results indicate gravida, parity, and spontaneous pregnancy losses all have strong-to-moderate positive correlations to each other, but not to any metabolites. This is not surprising, as spontaneous pregnancy losses contribute to the disparity between total pregnancies and births. We found that maternal age has a moderate positive correlation (r = 0.32) to unsaturated lipids and PC/PE (1656 cm^{-1}). Maternal plasma lipids have been correlated to the pathogenesis of PTB and shown to increase with maternal age.[37] Membrane phospholipids including PC generally increase throughout aging.[53] This implies advanced maternal age may correlate to a risk of PTB, and a lipidomic screening in the first trimester would be beneficial in patients >35 years of age. We found that patient BMI is moderately correlated to several metabolites, including an inverse strong to moderate correlation with carotenoids (1154 cm^{-1}, 1517 cm^{-1}) and PC/PE (1656 cm^{-1}). It has been shown in mouse models that dietary PC benefits individuals with diet-induced obesity, suggesting that diet may play a critical role in the risk of PTB for patients with a high BMI.[54] BMI also has a moderate joint correlation to AAs including tyrosine (1615 cm^{-1}), phenylalanine (1583 cm^{-1}), and tryptophan (1551 cm^{-1}). Obesity is associated with a reduced protein synthesis rate, resulting in increased plasma amino acid concentrations.[55] Further, BMI is also jointly correlated to NADH (1569 cm^{-1}); metabolic balance between white adipose tissue, which stores energy, and brown adipose tissue, which expends energy, is associated with the NAD+/NADH redox couple.[56] In addition, we found moderate-to-strong intra-metabolite correlations where metabolites that increase (Figure 2a–f) or decrease (Figure 2g–l) in PTB patients are jointly correlated. We also discovered that metabolites of the same class share a positive correlation, such as amino acids methionine (681 cm^{-1}), histidine (659 cm^{-1}), tyrosine (1615 cm^{-1}), and phenylalanine (1605 cm^{-1}). Finally, strong joint correlations between TCA cycle metabolites (citric acid, 940 cm^{-1}) and fatty acids (1173 cm^{-1}) and lipids/myristic acid (956 cm^{-1}) are also observed, demonstrating

that these metabolic pathways are intertwined in pregnancy and are critical to the onset of PTB.

In comparing our Raman findings to clinical measures, currently, no standard mechanisms exist to clinically predict PTB in the 1st trimester. Ultrasound measurement of cervical length (CL) is performed at ~16 weeks of gestation, but CL alone is neither an early predictor nor an accurate measure of risk of PTB. A recent meta-analysis of >1000 women showed that CL measurements at 18–22 weeks scored a poor AUC of 0.605 in predicting PTB, which slightly improved to 0.725 at 28–32 weeks of gestation.[57] They also concluded at 11–14 weeks gestation there were no differences in CL between those at risk of PTB and healthy pregnancies. Another prospective study of a low-risk population of 1113 women also concluded that CL measurement at 11–14 weeks of gestation does not show any predictive value (AUC = 0.55) for those that delivered <35 weeks gestation.[58] CL measurement is therefore sometimes combined with a fetal fibronectin (fFn) test, a protein marker in the cervicovaginal fluid, used in the 2nd trimester as a negative predictor for asymptomatic patients at high risk of PTB. In many clinical settings fFn is presented as a qualitative test with positive or negative result based on a concentration cut-off of 50 ng/ml.[59] But qualitative tests have poor accuracy reporting an AUC of 0.68.[60] Recent quantitative fFn tests have overall improved accuracy (AUC of 0.78) but these tests are not included in routine prenatal care. Patients are only tested when they present symptoms of PTB but do not show signs of cervical dilation. Therefore, current clinical measures are inadequate in early and accurate risk prediction of PTB, and innovative, cutting-edge technologies are imperative that fill this clinical gap in patient stratification.

Our compelling findings suggest that *(i)* metabolic profiling with Raman has the potential to address this critical clinical gap by examining changes in multiple metabolites in the 1st trimester that report a good predictive value (AUC of 0.839) for early assessment of those at risk of PTB. *(ii)* Tentative metabolic assignment from Raman combined with clinical and obstetric history of patients (Fig. 5c, d) scores an unprecedented AUC of 0.851 to enable potential risk assessment of PTB in the 1st trimester that surpasses current clinical standards. *(iii)* Raman-based screening also complements current prenatal care allowing a rapid (30 min measurement per sample), low-cost, and highly sensitive approach that requires minimal sample volumes (3 μL plasma). Therefore, Raman metabolic testing may be implemented as affordable maternal screening for patients as early as the first trimester to achieve a predictive risk scoring.

We compared Raman findings with standard protein assays of exploratory markers for PTB in the first trimester. Here, we performed ELISA (Figure 6) in n =11 healthy patients and n=10 PTB patients for four pro-inflammatory cytokines.[7, 61, 62] These include tumor necrosis factor-alpha (TNF-α), interleukin-12 (IL-12p20), interleukin-6 (IL-6), and granulocyte-macrophage colony-stimulating factor (GM-CSF). Three independent experiments for each cytokine and for each patient sample were performed i.e. each dot in Figure 6 represents the average of the experimental triplicate data for one patient. Values were normalized to total protein in the samples which were assessed with Bicinchoninic acid protein assay. Our results show no statistically significant differences in any of the cytokines when comparing patients that delivered preterm vs. those that delivered at term. A slight

decrease in GM-CSF in the PTB cohort is observed which corresponds well to findings in the literature,[63] but generally, first trimester cytokine levels did not have a clear correlation with an increased risk of PTB. These protein levels are not surprising, as cytokines and other proteins are already established in the literature to have low positive predictive values for accurate diagnosis.[64] This result is, in part, due to the poor sensitivity of ELISA and the inability to measure ultra-low levels of proteins. Further, such protein-based assays also have inter- and intra-assay discrepancies between studies due to variations in the type of sample studied and timing of sample collection.[7, 8] These challenges collectively make a comparative analysis of proteins difficult for PTB prediction and provide an opportunity to expand Raman metabolic profiling as an affordable screening platform for all pregnancies. tSNE of the ELISA results (Figure 6e) confirms that differences between PTB and healthy cohorts in the first trimester is inconclusive with these protein levels. tSNE parameters include an exaggeration value of 2, 3 PCA components, and a perplexity of 11.

3. Conclusions

In summary, this study demonstrates rapid, low-cost, and highly sensitive metabolic profiling with Raman in the first trimester plasma samples of pregnant patients. Raman data combined with unsupervised machine learning enable a predictive measure of patients early in the pregnancy who may be at risk of PTB. Tentative metabolite classes identified with Raman and validated with MS metabolomics showed an increase in lipids, fatty acids, membrane phospholipids, and metabolites of the TCA cycle; literature evidence shows these trends are expected in PTB pathogenesis. Based on the Raman peak assignments, we found that amino acids play a complex role, as some amino acids are necessary for placental growth and fetal development while others contribute to the inflammatory pathophysiology of PTB. We performed enrichment analysis to determine possible metabolic pathways that may be early predictors of PTB, and we found 18 significant pathways in Raman that also overlap with MS. Our analysis conservatively shows the relevance of Raman in metabolic assessment, which compliments the high specificity of MS in identifying a larger subset of metabolites. Finally, we demonstrate that the tentative metabolites identified with Raman, when combined with clinical information and obstetric history, enable a compelling screening tool for identifying patients who may be at risk of PTB in their first trimester. A correlation analysis with the Pearson coefficient shows that membrane phospholipids likely has a moderate correlation to maternal age and BMI, indicating that a lipidomic screening in the first trimester would benefit patients of advanced maternal age. BMI also correlated to multiple amino acids, indicating that patients with high BMI may be offered additional recommendations for diet management or medications. Our ELISA analysis of key pro-inflammatory cytokines showed that first trimester cytokine levels did not have a clear relationship to an increased risk of PTB, rendering it difficult to predict the mechanisms of PTB based on inflammatory proteins alone. We note that in this study, samples were obtained from an academic biobank where patient plasma (and other biofluids) are continuously banked for research and education; therefore, patients were not targeted or recruited specifically for this study. The purpose of our work was to demonstrate the capabilities of Raman in early diagnosis of PTB and we expect that our results may be improved and more metabolites may be identified in a larger, targeted cohort study in

future. We envision that Raman spectral measurements of maternal biofluids combined with machine learning has the potential to complement current prenatal care, offering a predictive tool that may be useful for risk stratification of PTB. Our study is also expected to advance maternal and fetal health in rural areas enabling affordable screening for patients throughout their pregnancy.

4. Experimental Methods

Materials:

Ribitol, nonadecanoic acid, methoxamine HCl, ethanol, acetone and BSTFA + 1% TMCS were obtained from Sigma-Aldrich, Inc. (St. Louis, Missouri, USA). Methanol, water, chloroform, and hexane (HPLC grade) were obtained from Fisher Chemical; Thermo Fisher Scientific, Inc. (Waltham, MA, USA).

Human plasma samples:

All patients provided informed consent to have blood collected by the Perinatal Family Tissue Bank from the University of Iowa Department of Obstetrics and Gynecology (IRB#200910784). Blood was collected in ACD-A tubes (Becton Dickinson), centrifuged, aliquoted, snap-frozen, and stored at −80 °C to isolate the plasma. All samples and paired clinical data received from the tissue bank were de-identified to maintain patients' anonymity.

Raman mapping and analysis:

A Renishaw inVia Raman confocal microscope with a 785 nm laser was used to take scans of the plasma with WiRE 5.4 software. A 120 mW laser was operated at 100% power with 1200 lines per mm grating. A spot size of 1 μm resolution and a 50 × objective were used to obtain spectra. Static scans were taken from 600 to 1700 cm^{-1} with 10 sec of exposure time and one accumulation. The instrument was calibrated using an external silicon standard wafer at 520.5 nm daily, and line maps were used to take 100 spectra per patient. The plasma samples were prepared by thawing and aliquoting 3 μL on 2 mm × 2 mm CaF_2 disks from Crystran Ltd followed by drying in a 37 °C oven for 20 min prior to measurement. Smoothing, baseline correction, cosmic ray removal, and standard normal variate (SNV) method normalization were applied with a custom MATLAB code. The spectra were smoothed using a Savitzky and Golay filter with a 10^{th} order polynomial and coefficient value of 45 points.[65] The fluorescent background was removed using a modified polyfit method with a 9th order polynomial fit with a threshold value of 0.0001.[66] To visualize data clustering, the MATLAB Statistics and Machine Learning Toolbox implementation of t-distributed stochastic neighbor embedding (tSNE) was used via a custom MATLAB code (R2021b). This method is a nonlinear, unsupervised, multivariate dimensionality reduction technique. Parameters included the Exact algorithm with a Cosine distance function, an exaggeration value of two, and PCA components and perplexity which varied by figure and are included in the text.

Mass spectrometry Metabolomics:

Plasma samples were submitted to the W.M. Keck Metabolomics Research Laboratory (Office of Biotechnology, Iowa State University, Ames IA) for non-targeted metabolomic analysis. Sample preparation was conducted using a modified methanolic extraction and sample preparation methods established by Moritz and co-workers.[67] 50 μL of each sample were spiked with internal standards (10 μg of nonadecanoic acid (1 mg/mL in methanol) (Sigma-Aldrich CO., St. Louis, MO), 10 μg ribitol (1 mg/mL in water) (Sigma-Aldrich CO., St. Louis, MO) and the extraction initiated with the addition of 0.9 mL of 80% ice-cold LC-MS grade methanol with 20% LC-MS grade water (Fisher Scientific, Waltham, MA). Samples were vortexed for 90 sec, rested on ice for 10 min, followed by more vortexing for 90 sec then placed into an ice-cold sonication water bath for 10 min at full output power. Samples were then vortexed again for 90 sec and centrifuged for 7 min at 13,000 × g at room temperature. The supernatants were recovered, and the remaining insoluble pellets were re-extracted with additional volume of 0.9 mL of 80% ice-cold methanol, then the supernatant extracts were pooled.

Six hundred microliters of the combined extracts were dried using a speed-vac concentrator for 10 h prior to derivation.[68] Samples were derivatized with 50 μL of methoxyamine hydrochloride (20 mg/mL in pyridine) initially added to the dried extracts followed by a 1.5 h incubation at 30 °C. Subsequently trimethylsilylation (TMS) was performed by addition of 70 μL of bis-trimethyl silyl trifluoroacetamide with 1% Trimethylchlorosilane (BSTFA + 1% TMCS) for 30 min at 60 °C. GC-MS characterized the derivatized samples at the ISU W.M. Keck Metabolomics Research Laboratory. GC-MS analyses were performed with an Agilent 6890 gas chromatograph coupled to a model 5973 Mass Selective Detector (Agilent Technologies, Santa Clara, CA). The column used was HP-5MSI 5% Phenyl Methyl Silox with 30 m × 250 μM × 0.25 μm film thickness (Agilent Technologies). One microliter of the sample was injected with the inlet operating in splitless mode and held at a constant temperature of 280 °C. The oven temperature was programmed as follows: an initial temperature of 70 °C was increased to 250 °C at 15 °C/min, followed by an increase of 20 °C/min until reaching 320 °C, which was held for 5 min. Helium was used as a carrier gas at a flow rate of 1 mL/min. The MS transfer line was held at 280 °C. Mass Spectrometry detection was performed using electron ionization at 70 eV and source temperature and quadrupole temperature were set at 230 °C and 150 °C, respectively. The mass data was collected in the range from m/z 40 to m/z 800. Identification and quantification were conducted using AMDIS (Automated Mass spectral Deconvolution and Identification System, National Institute of Standards and Technology (Gaithersburg, MD)) with a manually curated retention indexed GC-MS library with additional identification performed using the NIST17 and Wiley 11 GC-MS spectral library (Agilent Technologies, Santa Clara, CA). Final quantification was calculated by integrating the corresponding peak areas relative to the area of the internal standards. Raw data was normalized to the amount of tissue used. Statistical evaluation of the non-targeted GC-MS data was conducted with the R-based statistical package MetaboAnalyst.[69] Important molecular features were elucidated using the MetaboAnalyst multivariate analysis tools.

The supernatant extracts were pooled and filtered with 0.2 μm centrifugal filters (Cat. No. UFC30LG25, Millipore Sigma, Burlington, MA) and subjected to LC-MS analysis. LC separations were performed with an Agilent Technologies 1290 Infinity Binary Pump UHPLC instrument equipped with an Agilent Technologies Eclipse C18 1.8 μm 2.1 mm × 100 mm analytical column that was coupled to an Agilent Technologies 6540 UHD Accurate-Mass Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA). Six μL of each sample were injected into the LC system. Chromatography was carried out at 40 °C with a flow rate of 0.400 mL/min. Running solvents were A: water with 5mM ammonium formate and 0.1% formic acid; B: 25% iso-propanol in acetonitrile with 5 mM ammonium formate and 0.1% formic acid. Initial solvent conditions were 0% B which increased on a linear gradient to 100% B over 15 min, 100% B was held for 5 min before returning to 0% B over 2 min. After each LC-MS acquisition, a 6-min post run at 0% B was conducted. Metabolites were detected using electrospray ionization in both negative and positive ionization modes. Nitrogen was used as the service gas for the ion source with a drying gas flow rate of 12 L/min at a temperature of 350 °C, a nebulizing pressure of 25 psi, and a sheath gas flow of 11 L/min at 400 °C. The capillary and nozzle voltages were 4000 and 1750 volts respectively. The mass spectrometer was operated in high resolution (4 Gz) mode with a scan range from m/z 100 to m/z 1700. An acquisition rate of 1.5 spectra per sec was used. Reference masses were monitored for continuous mass calibration during LC-MS data acquisition: m/z 121.050873 with m/z 922.009698 for positive mode and m/z 112.985587 with 1033.988109 for negative mode. Data evaluation and peak detection were performed using Agilent MassHunter Qualitative Analysis (version 10.0) and Mass Profiler programs (version 8.0) (Agilent Technologies, Santa Clara, CA). Metabolite peaks were identified using accurate mass spectral analysis compared to the METLIN database.[70]

Enzyme-linked immunosorbent assay (ELISA):

IL-6 (88-7066-88), IL-12p20 (88-7126-88), GM-CSF (88-7346-88), and TNF-α (88-8337-88) ELISA kits for humans were purchased from Thermo Fisher (Waltham, MA). Corning® 96 well half-area microplates were used for ELISA analysis. Total protein was measured by BCA protein assay (Thermo Fisher 23225). The assays were performed according to the manufacturer's protocol. GraphPad Prism 8.2.1 software was used to analyze the data.

MetaboAnalyst Methods:

MetaboAnalyst, a web based interface for analysis of metabolic data, was used to generate figures for AUC-ROC analysis (Figure 1d), correlation analysis (Figure 2m), MS and RS metabolites intensity heat maps with auto-scaled features per row (Figure 3 and S4), enrichment analysis (Figure 4a,b) and Pearson correlation heat map from normalized Raman intensity values (Figure 5e). The Statistical Analysis [one factor] module was used to create the heat maps and correlation analysis for the Raman and MS data, and [metadata table] created the correlation bar graph. Heat maps were generated using the default MetaboAnalyst parameters: normalized data-to-scalable features, Euclidean distance, Ward clustering, and analysis with Student's *t*-test. The Enrichment Analysis module created the bar graphs detailing metabolic pathways for Raman and MS data. For Raman enrichment analysis, the individual Raman peak assignments (corresponding to different

metabolites) and the peak intensities for all of the patients in this study were uploaded into Metaboanalyst. Here the intensity of the Raman peaks was used as a "proxy" for metabolite concentrations to obtain the corresponding metabolic pathways. Generic peaks in Raman such as carbohydrates (1017 cm^{-1}), amino acids (1206 cm^{-1}), amide III (1243 cm^{-1}), amide I (1672 cm^{-1}) and DNA (805, 1420 cm^{-1}) were not used in the enrichment analyses since they are representative of a broad family of metabolites. The enrichment analyses provides p-values for the most probable metabolic pathways. But note that the enrichment p-values provided by Metaboanalyst are different than the p-values obtained from statistical analysis. The enrichment p-value is defined as the probability of obtaining n or more pathways that give rise to a cumulative hypergeometric distribution. Metaboanalyst selects the hypergeometric test for over-representation analysis (ORA). ORA is a statistical method that determines whether metabolites from a pre-defined set (such as the KEGG database) are present more than would be expected (over-represented) in the data we upload. MS enrichment pathways was obtained similarly. Quantified enrichment analysis was used with the ID type as compound names and feature type as metabolites with a categorical group label. The KEGG pathway was used to identify metabolic pathways in the Raman and MS data. Finally, the Biomarker Analysis module was utilized to prepare AUC-ROC curves in Tester mode. The peak values were uploaded for every sample in both cohorts. Only the strongest peak for each metabolite was used and additional peaks belonging to the same metabolite was removed from the analysis as only a single can be used in Metaboanalyst. Metaboanalyst offers the option to compute and add statistically significant ratios to this analysis. But we chose to not add this option to avoid overfitting our model. This option can be applied when larger sample numbers are available.

Statistical analysis:

Data is presented as standard error of the mean. Unpaired, two-sided, homoscedastic Student's t-tests were used in excel to generate p values. P values of < 0.05 and < 0.01 are indicated as * and **.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

LS acknowledges support from the National Institutes of Health (NIH) award R21HD100685-01. SU and GC acknowledges support from the NIH R01EB029756-01A1. CYL and ECL acknowledge support from the Ministry of Science and Technology of Taiwan (MOST) award 109-2113-M-194-010-MY3. XW acknowledges support from the congressionally directed medical research program (CDMRP) award W81XWH-18-1-0139. MS and DS acknowledge support from the NIH R01HD089940, NIH 3UL1TR002537-02S1, NIH UL1TR002537, and American Heart Association (AHA) 15SFRN23480000 awards. RB acknowledges support from the NIH R21HD100685-01, NIH R01EB029756-01A1, and CDMRP W81XWH-20-1-0620 awards.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

(1). Walani SR Global Burden of Preterm Birth. Int J Gynaecol Obstet 2020, 150 (1), 31–33. DOI: 10.1002/ijgo.13195. [PubMed: 32524596]

(2). Tekesin I; Eberhart LH; Schaefer V; Wallwiener D; Schmidt S Evaluation and Validation of a New Risk Score (Cleopatra Score) to Predict the Probability of Premature Delivery for Patients with Threatened Preterm Labor. Ultrasound Obstet Gynecol 2005, 26 (7), 699–706. DOI: 10.1002/uog.2633. [PubMed: 16308893]

(3). Goffinet F Primary Predictors of Preterm Labour. BJOG 2005, 112 (1), 38–47.

(4). Berghella V; Saccone G Cervical Assessment by Ultrasound for Preventing Preterm Delivery. Cochrane Database Syst Rev 2019, 9, CD007235. DOI: 10.1002/14651858.CD007235.pub4. [PubMed: 31553800]

(5). Soto EE; Hernandez-Andrade E; Huntley ES; Blackwell SC Implementing a Universal Cervical Length Screening Program in a Large Hospital System: It Takes Some Time to Achieve Consistent Results. Gynecol Obstet Invest 2022, 87 (2), 124–132. DOI: 10.1159/000524361. [PubMed: 35354147]

(6). Gibson JL; Macara LM; Owen P; Young D; Macauley J; Mackenzie F Prediction of Preterm Delivery in Twin Pregnancy: A Prospective, Observational Study of Cervical Length and Fetal Fibronectin Testing. Ultrasound Obstet Gynecol 2004, 23 (6), 561–566. DOI: 10.1002/uog.1048. [PubMed: 15170796]

(7). Menon R; Torloni MR; Voltolini C; Torricelli M; Merialdi M; Betrán AP; Widmer M; Allen T; Davydova I; Khodjaeva Z; Thorsen P; Kacerovsky M; Tambor V; Massinen T; Nace J; Arora C Biomarkers of Spontaneous Preterm Birth: An Overview of the Literature in the Last Four Decades. Reprod Sci 2011, 18 (11), 1046–1070. DOI: 10.1177/1933719111415548. [PubMed: 22031189]

(8). Polettini J; Cobo T; Kacerovsky M; Vinturache AE; Laudanski P; Peelen MJ; Helmer H; Lamont RF; Takeda J; Lapointe J; Torloni MR; Zhong N; Menon R Biomarkers of Spontaneous Preterm Birth: A Systematic Review of Studies Using Multiplex Analysis. J Perinat Med 2017, 45 (1), 71–84. DOI: 10.1515/jpm-2016-0097. [PubMed: 27514075]

(9). Tsiartas P; Holst RM; Wennerholm UB; Hagberg H; Hougaard DM; Skogstrand K; Pearce BD; Thorsen P; Kacerovsky M; Jacobsson B Prediction of Spontaneous Preterm Delivery in Women with Threatened Preterm Labour: A Prospective Cohort Study of Multiple Proteins in Maternal Serum. BJOG 2012, 119 (7), 866–873. DOI: 10.1111/j.1471-0528.2012.03328.x. [PubMed: 22530716]

(10). Liang L; Rasmussen MH; Piening B; Shen X; Chen S; Röst H; Snyder JK; Tibshirani R; Skotte L; Lee NC; Contrepois K; Feenstra B; Zackriah H; Snyder M; Melbye M Metabolic Dynamics and Prediction of Gestational Age and Time to Delivery in Pregnant Women. Cell 2020, 181 (7), 1680–1692.e1615. DOI: 10.1016/j.cell.2020.05.002. [PubMed: 32589958]

(11). Bunaciu AA; Aboul-Enein HY; Fleschin Ş Vibrational Spectroscopy in Clinical Analysis. Applied Spectroscopy Reviews 2015, 50 (2), 176–191.

(12). Wen X; Ou YC; Bogatcheva G; Thomas G; Mahadevan-Jansen A; Singh B; Lin EC; Bardhan R Probing Metabolic Alterations in Breast Cancer in Response to Molecular Inhibitors with Raman Spectroscopy and Validated with Mass Spectrometry. Chem Sci 2020, 11 (36), 9863–9874. DOI: 10.1039/d0sc02221g. [PubMed: 34094246]

(13). Ou Y-C; Wen X; Johnson CA; Shae D; Ayala O; Webb JA; Lin EC; DeLapp RC; Boyd KL; Richmond A; Mahadevan-Jansen A; Rafat M; Wilson JT; Balko JM; N. TM; Vilgelm AE; Bardhan R Multimodal Multiplexed Immunoimaging with Nanostars to Detect Multiple Immunomarkers and Monitor Response to Immunotherapies. ACS Nano 2020, 14 (1), 651–663. [PubMed: 31851488]

(14). Ou Y-C; Wen X; Bardhan R Cancer Immunoimaging with Smart Nanoparticles. Trends in biotechnology 2019, 38 (4), 388–403. [PubMed: 31812371]

(15). Wen X; Ou Y-C; Zarick HF; Zhang X; Hmelo AB; Victor Q; Paul EP; Slocik JM; Naik RR; Bellan LM; Lin EC; Bardhan R Prada: Portable Reusable Accurate Diagnostics with Nanostar Antennas for Multiplexed Biomarker Screening. Bioengineering and Translational Medicine 2020, 5 (3), e10165. DOI: 10.1002/btm2.10165. [PubMed: 33005736]

(16). Talari ACS; Movasaghi Z; Rehman S; Rehman IU Raman Spectroscopy of Biological Tissues. Appl Spectrosc Rev 2015, 50 (1), 46–111. DOI: 10.1080/05704928.2014.923902.

(17). Zhu G; Zhu X; Fan Q; Wan X Raman Spectra of Amino Acids and Their Aqueous Solutions. Spectrochim Acta A 2011, 78 (3), 1187–1195. DOI: 10.1016/j.saa.2010.12.079.

(18). De Gelder J; De Gussem K; Vandenabeele P; Moens L Reference Database of Raman Spectra of Biological Molecules. J. Raman Spectrosc 2007, 38 (9), 1133–1147.

(19). Czamara K; Majzner K; Pacia MZ; Kochan K; Kaczor A; Baranska M Raman Spectroscopy of Lipids: A Review. J. Raman Spectrosc 2015, 46 (1), 4–20. DOI: 10.1002/jrs.4607.

(20). Ricciardi A; Piuri G; Della Porta M; Mazzucchelli S; Bonizzi A; Truffi M; Seviere M; Allevi R; Corsi F; Cazzola R; Morasso C Raman Spectroscopy Characterization of the Major Classes of Plasma Lipoproteins. Vibrational Spectroscopy 2020, 109, 103073.

(21). Atkins CG; Buckley K; Blades MW; Turner RF Raman Spectroscopy of Blood and Blood Components. Applied spectroscopy 2017, 71 (5), 767–793. [PubMed: 28398071]

(22). van der Maaten L; Hinton G Visualizing Data Using T-Sne. Journal of Machine Learning Research 2008, 9 (11), 2579–2605.

(23). Hu S; Li H; Chen C; Chen C; Zhao D; Dong B; Lv X; Zhang K; Xie Y Raman Spectroscopy Combined with Machine Learning Algorithms to Detect Adulterated Suichang Native Honey. Sci Rep 2022, 12 (1), 3456. [PubMed: 35236873]

(24). Ullah R; Khan S; Javaid S; Ali H; Bilal M; Saleem M Raman Spectroscopy Combined with a Support Vector Machine for Differentiating between Feeding Male and Female Infants Mother's Milk. Biomed Opt Express 2018, 9 (2), 844–851. DOI: 10.1364/BOE.9.000844. [PubMed: 29552417]

(25). He C; Zhu S; Wu X; Zhou J; Chen Y; Qian X; Ye J Accurate Tumor Subtype Detection with Raman Spectroscopy Via Variational Autoencoder and Machine Learning. ACS Omega 2022, 7 (12), 10458–10468. DOI: 10.1021/acsomega.1c07263 From NLM PubMed-not-MEDLINE. [PubMed: 35382336]

(26). Menon R; Conneely KN; Smith AK DNA Methylation: An Epigenetic Risk Factor in Preterm Birth. Reprod Sci 2012, 19 (1), 6–13. DOI: 10.1177/1933719111424446. [PubMed: 22228737]

(27). Park B; Khanam R; Vinayachandran V; Baqui AH; London SJ; Biswal S Epigenetic Biomarkers and Preterm Birth. Environmental Epigenetics 2020, 6 (1), 5. DOI: 10.1093/eep/dvaa005.

(28). Daum R; Brauchle EM; Berrio DAC; Jurkowski TP; Schenke-Layland K Non-Invasive Detection of DNA Methylation States in Carcinoma and Pluripotent Stem Cells Using Raman Microspectroscopy and Imaging. Sci. Rep 2019, 9 (1), 1–13. [PubMed: 30626917]

(29). Gomez-Lopez N; Romero R; Schwenkel G; Garcia-Flores V; Panaitescu B; Varrey A; Ayoub F; Hassan SS; Phillippe M Cell-Free Fetal DNA Increases Prior to Labor at Term and in a Subset of Preterm Births. Reprod Sci 2020, 27 (1), 218–232. DOI: 10.1007/s43032-019-00023-6. [PubMed: 32046392]

(30). Darghahi R; Mobaraki-Asl N; Ghavami Z; Pourfarzi F; Hosseini-Asl S; Jalilvand F Effect of Cell-Free Fetal DNA on Spontaneous Preterm Labor. J Adv Pharm Technol Res 2019, 10 (3), 117–120. DOI: 10.4103/japtr.JAPTR_371_18. [PubMed: 31334093]

(31). Lizewska B; Teul J; Kuc P; Lemancewicz A; Charkiewicz K; Goscik J; Kacerovsky M; Menon R; Miltyk W; Laudanski P Maternal Plasma Metabolomic Profiles in Spontaneous Preterm Birth: Preliminary Results. Mediators Inflamm 2018, 2018, 9362820. DOI: 10.1155/2018/9362820. [PubMed: 29670470]

(32). Elshenawy S; Pinney SE; Stuart T; Doulias P-T; Zura G; Parry S; Elovitz MA; Bennett MJ; Bansal A; Strauss III JF The Metabolomic Signature of the Placenta in Spontaneous Preterm Birth. Int. J. Mol. Sci 2020, 21 (3), 1043. [PubMed: 32033212]

(33). Elshenawy S; Pinney SE; Stuart T; Doulias P-T; Zura G; Parry S; Elovitz MA; Bennett MJ; Bansal A; Strauss III JF; Ischiropoulos H; Simmons RA The Metabolomic Signature of the Placenta in Spontaneous Preterm Birth. International Journal of Molecular Sciences 2020, 21 (3), 1043. [PubMed: 32033212]

(34). Fiedor J; Burda K Potential Role of Carotenoids as Antioxidants in Human Health and Disease. Nutrients 2014, 6 (2), 466–488. DOI: 10.3390/nu6020466. [PubMed: 24473231]

(35). You YA; Hwang SY; Kim SM; Park S; Lee GI; Ansari A; Lee J; Kwon Y; Kim YJ Identification of Indicators for Preterm Birth Using Retinoid Metabolites. Metabolites 2021, 11 (7), 443. DOI: 10.3390/metabo11070443. [PubMed: 34357337]

(36). Ghazvini S; Uthaman S; Synan L; Lin, E. C; Sarkar S; Santillan MK; Santillan DA; Bardhan R Predicting the Onset of Preeclampsia by Longitudinal Monitoring of Metabolic Changes Throughout Pregnancy with Raman Spectroscopy. Biotechnology & translational medicine (under review).

(37). Chen Y; He B; Liu Y; Aung MT; Rosario-Pabón Z; Vélez-Vega CM; Alshawabkeh A; Cordero JF; Meeker JD; Garmire LX Maternal Plasma Lipids Are Involved in the Pathogenesis of Preterm Birth. Gigascience 2022, 11, 4. DOI: 10.1093/gigascience/giac004.

(38). Kurian NK; Modi D Extracellular Vesicle Mediated Embryo-Endometrial Cross Talk During Implantation and in Pregnancy. J Assist Reprod Genet 2019, 36 (2), 189–198. DOI: 10.1007/s10815-018-1343-x. [PubMed: 30362057]

(39). Aung MT; Ashrap P; Watkins DJ; Mukherjee B; Rosario Z; Vélez-Vega CM; Alshawabkeh AN; Cordero JF; Meeker JD Maternal Lipidomic Signatures in Relation to Spontaneous Preterm Birth and Large-for-Gestational Age Neonates. Sci Rep 2021, 11 (1), 8115. DOI: 10.1038/s41598-021-87472-9. [PubMed: 33854141]

(40). Laudanski P; Charkiewicz K; Kisielewski R; Kuc P; Koc-Zorawska E; Raba G; Kraczkowski J; Dymicka-Piekarska V; Chabowski A; Kacerovsky M; Jacobsson B; Zabielski P; Blachnio-Zabielska A Plasma C16-Cer Levels Are Increased in Patients with Preterm Labor. Prostaglandins & Other Lipid Mediators 2016, 123, 40–45. [PubMed: 27184754]

(41). Holm MB; Bastani NE; Holme AM; Zucknick M; Jansson T; Refsum H; Morkrid L; Blomhoff R; Henriksen T; Michelsen TM Uptake and Release of Amino Acids in the Fetal-Placental Unit in Human Pregnancies. PLoS One 2017, 12 (10), e0185760. DOI: 10.1371/journal.pone.0185760. [PubMed: 28982184]

(42). Karahoda R Dynamics of Tryptophan Metabolic Pathways in Human Placenta and Placental-Derived Cells: Effect of Gestation Age and Trophoblast Differentiation. Frontiers 2020, 8, 574034. DOI: 10.3389/fcell.2020.574034.

(43). De Groof F; Huang L; van Vliet I; Voortman GJ; Schierbeek H; Roksnoer LC; Vermes A; Chen C; Huang Y; van Goudoever JB Branched-Chain Amino Acid Requirements for Enterally Fed Term Neonates in the First Month of Life. Am J Clin Nutr 2014, 99 (1), 62–70. DOI: 10.3945/ajcn.112.038927. [PubMed: 24284437]

(44). Jehan F; Sazawal S; Baqui AH; Nisar MI; Dhingra U; Khanam R; Ilyas M; Dutta A; Mitra DK; Mehmood U; Deb S; Mahmud A; Hotwani A; Ali SM; Rahman S; Nizar A; Ame SM; Moin MI; Muhammad S; Chauhan A; Begum N; Khan W; Das S; Ahmed S; Hasan T; Khalid J; Rizvi SJR; Juma MH; Chowdhury NH; Kabir F; Aftab F; Quaiyum A; Manu A; Yoshida S; Bahl R; Rahman A; Pervin J; Winston J; Musonda P; Stringer JSA; Litch JA; Ghaemi MS; Moufarrej MN; Contrepois K; Chen S; Stelzer IA; Stanley N; Chang AL; Hammad GB; Wong RJ; Liu C; Quaintance CC; Culos A; Espinosa C; Xenochristou M; Becker M; Fallahzadeh R; Ganio E; Tsai AS; Gaudilliere D; Tsai ES; Han X; Ando K; Tingle M; Maric I; Wise PH; Winn VD; Druzin ML; Gibbs RS; Darmstadt GL; Murray JC; Shaw GM; Stevenson DK; Snyder MP; Quake SR; Angst MS; Gaudilliere B; Aghaeepour N; Alliance for Maternal and Newborn Health Improvement, t. G. A. t. P. P. a. S., and the Prematurity Research Center at Stanford University. Multiomics Characterization of Preterm Birth in Low- and Middle-Income Countries. JAMA Netw Open 2020, 3 (12), e2029655. DOI: 10.1001/jamanetworkopen.2020.29655. [PubMed: 33337494]

(45). Movasaghi Z Raman Spectroscopy of Biological Tissues. Applied Spectroscopy Reviews 2007, 42, 493–541.

(46). Shi JW; Lai ZZ; Yang HL; Yang SL; Wang CJ; Ao D; Ruan LY; Shen HH; Zhou WJ; Mei J; Fu Q; Li MQ Collagen at the Maternal-Fetal Interface in Human Pregnancy. Int J Biol Sci 2020, 16 (12), 2220–2234. DOI: 10.7150/ijbs.45586. [PubMed: 32549767]

(47). Maddahi MS; Dolatian M; Khoramabadi M; Talebi A Correlation of Maternal-Fetal Attachment and Health Practices During Pregnancy with Neonatal Outcomes. Electron Physician 2016, 8 (7), 2639–2644. DOI: 10.19082/2639. [PubMed: 27648191]

(48). Lopes KR; Souza ASR; Figueiroa JN; Alves JGB Correlation between Pre-Pregnancy Body Mass Index and Maternal Visceral Adiposity with Fetal Biometry During the Second Trimester. Int J Gynaecol Obstet 2017, 138 (2), 133–137. [PubMed: 28485827]

(49). Cuny H; Rapadas M; Gereis J; Martin EM; Kirk RB; Shi H; Dunwoodie SL Nad Deficiency Due to Environmental Factors or Gene–Environment Interactions Causes Congenital Malformations and Miscarriage in Mice. Proceedings of the National Academy of Sciences 2020, 117 (7), 3738–3747.

(50). Czumaj A; Szrok-Jurga S; Hebanowska A; Turyn J; Swierczynski J; Sledzinski T; Stelmanska E The Pathophysiological Role of Coa. Int J Mol Sci 2020, 21 (23), 9057. DOI: 10.3390/ijms21239057. [PubMed: 33260564]

(51). Fukuwatari T; Shibata K Nutritional Aspect of Tryptophan Metabolism. Int J Tryptophan Res 2013, 6 (Suppl 1), 3–8. DOI: 10.4137/IJTR.S11588. [PubMed: 23922498]

(52). Wang P; Chen H; Tian J; Dai Z; Zou X Electrochemical Evaluation of DNA Methylation Level Based on the Stoichiometric Relationship between Purine and Pyrimidine Bases. Biosens Bioelectron 2013, 45, 34–39. DOI: 10.1016/j.bios.2013.01.057. [PubMed: 23454340]

(53). Dorninger F; Moser AB; Kou J; Wiesinger C; Forss-Petter S; Gleiss A; Hinterberger M; Jungwirth S; Fischer P; Berger J Alterations in the Plasma Levels of Specific Choline Phospholipids in Alzheimer's Disease Mimic Accelerated Aging. J Alzheimers Dis 2018, 62 (2), 841–854. DOI: 10.3233/JAD-171036. [PubMed: 29480199]

(54). Lee HS; Nam Y; Chung YH; Kim HR; Park ES; Chung SJ; Kim JH; Sohn UD; Kim H-C; Oh KW; Jeong JH Beneficial Effects of Phosphatidylcholine on High-Fat Diet-Induced Obesity, Hyperlipidemia and Fatty Liver in Mice. Life Sciences 2014, 118 (1), 7–14. [PubMed: 25445436]

(55). Simonson M; Boirie Y; Guillet C Protein, Amino Acids and Obesity Treatment. Rev Endocr Metab Disord 2020, 21 (3), 341–353. DOI: 10.1007/s11154-020-09574-5. [PubMed: 32827096]

(56). Luo C; Yang C; Wang X; Chen Y; Liu X; Deng H Nicotinamide Reprograms Adipose Cellular Metabolism and Increases Mitochondrial Biogenesis to Ameliorate Obesity. J. Nutr. Biochem 2022, 107, 109056. [PubMed: 35609856]

(57). Thain S; Yeo GSH; Kwek K; Chern B; Tan KH Spontaneous Preterm Birth and Cervical Length in a Pregnant Asian Population. PLoS One 2020, 15 (4), e0230125. DOI: 10.1371/journal.pone.0230125. [PubMed: 32282819]

(58). Antsaklis P; Daskalakis G; Pilalis A; Papantoniou N; Mesogitis S; Antsaklis A The Role of Cervical Length Measurement at 11–14 Weeks for the Prediction of Preterm Delivery. J Matern Fetal Neonatal Med 2011, 24 (3), 465–470. DOI: 10.3109/14767058.2010.501124. [PubMed: 20608797]

(59). Hezelgrave NL; Shennan AH Quantitative Fetal Fibronectin to Predict Spontaneous Preterm Birth: A Review. Womens Health (Lond) 2016, 12 (1), 121–128. DOI: 10.2217/whe.15.74. [PubMed: 26652920]

(60). Abbott DS; Hezelgrave NL; Seed PT; Norman JE; David AL; Bennett PR; Girling JC; Chandirimani M; Stock SJ; Carter J; Cate R; Kurtzman J; Tribe RM; Shennan AH Quantitative Fetal Fibronectin to Predict Preterm Birth in Asymptomatic Women at High Risk. Obstet Gynecol 2015, 125 (5), 1168–1176. DOI: 10.1097/AOG.0000000000000754. [PubMed: 25932845]

(61). Vogel I; Goepfert AR; Thorsen P; Skogstrand K; Hougaard DM; Curry AH; Cliver S; Andrews WW Early Second-Trimester Inflammatory Markers and Short Cervical Length and the Risk of Recurrent Preterm Birth. J Reprod Immunol 2007, 75 (2), 133–140. DOI: 10.1016/j.jri.2007.02.008. [PubMed: 17442403]

(62). Ferguson KK; McElrath TF; Chen YH; Mukherjee B; Meeker JD Longitudinal Profiling of Inflammatory Cytokines and C-Reactive Protein During Uncomplicated and Preterm Pregnancy. Am J Reprod Immunol 2014, 72 (3), 326–336. DOI: 10.1111/aji.12265. [PubMed: 24807462]

(63). Curry AE; Thorsen P; Drews C; Schendel D; Skogstrand K; Flanders WD; Hougaard D; Olsen J; Vogel I First-Trimester Maternal Plasma Cytokine Levels, Pre-Pregnancy Body Mass Index, and Spontaneous Preterm Delivery. Acta Obstet Gynecol Scand 2009, 88 (3), 332–342. DOI: 10.1080/00016340802702219. [PubMed: 19241227]

(64). Lucaroni F; Morciano L; Rizzo G; D' Antonio F; Buonuomo E; Palombi L; Arduini D Biomarkers for Predicting Spontaneous Preterm Birth: An Umbrella Systematic Review. J Matern Fetal Neonatal Med 2018, 31 (6), 726–734. DOI: 10.1080/14767058.2017.1297404. [PubMed: 28274163]

(65). Huang J; Shi T; Gong B; Li X; Liao G; Tang Z Fitting an Optical Fiber Background with a Weighted Savitzky-Golay Smoothing Filter for Raman Spectroscopy. Appl Spectrosc 2018, 72 (11), 1632–1644. DOI: 10.1177/0003702818785884. [PubMed: 30109810]

(66). Lieber CA; Mahadevan-Jansen A Automated Method for Subtraction of Fluorescence from Biological Raman Spectra. Applied spectroscopy 2003, 57 (11), 1363–1367. [PubMed: 14658149]

(67). A J; Trygg J; Gullberg J; Johansson AI; Jonsson P; Antti H; Marklund SL; Moritz T Extraction and Gc/Ms Analysis of the Human Blood Plasma Metabolome. Anal Chem 2005, 77 (24), 8086–8094. DOI: 10.1021/ac051211v. [PubMed: 16351159]

(68). Koek MM; Muilwijk B; van der Werf MJ; Hankemeier T Microbial Metabolomics with Gas Chromatography/Mass Spectrometry. Anal Chem 2006, 78 (4), 1272–1281. DOI: 10.1021/ac051683+. [PubMed: 16478122]

(69). Pang Z; Chong J; Zhou G; de Lima Morais DA; Chang L; Barrette M; Gauthier C; Jacques P; Li S; Xia J Metaboanalyst 5.0: Narrowing the Gap between Raw Spectra and Functional Insights. Nucleic Acids Res 2021, 49 (W1), W388–W396. DOI: 10.1093/nar/gkab382. [PubMed: 34019663]

(70). Baker ES; Patti GJ Perspectives on Data Analysis in Metabolomics: Points of Agreement and Disagreement from the 2018 Asms Fall Workshop. J Am Soc Mass Spectrom 2019, 30 (10), 2031–2036. DOI: 10.1007/s13361-019-02295-3. [PubMed: 31440979]

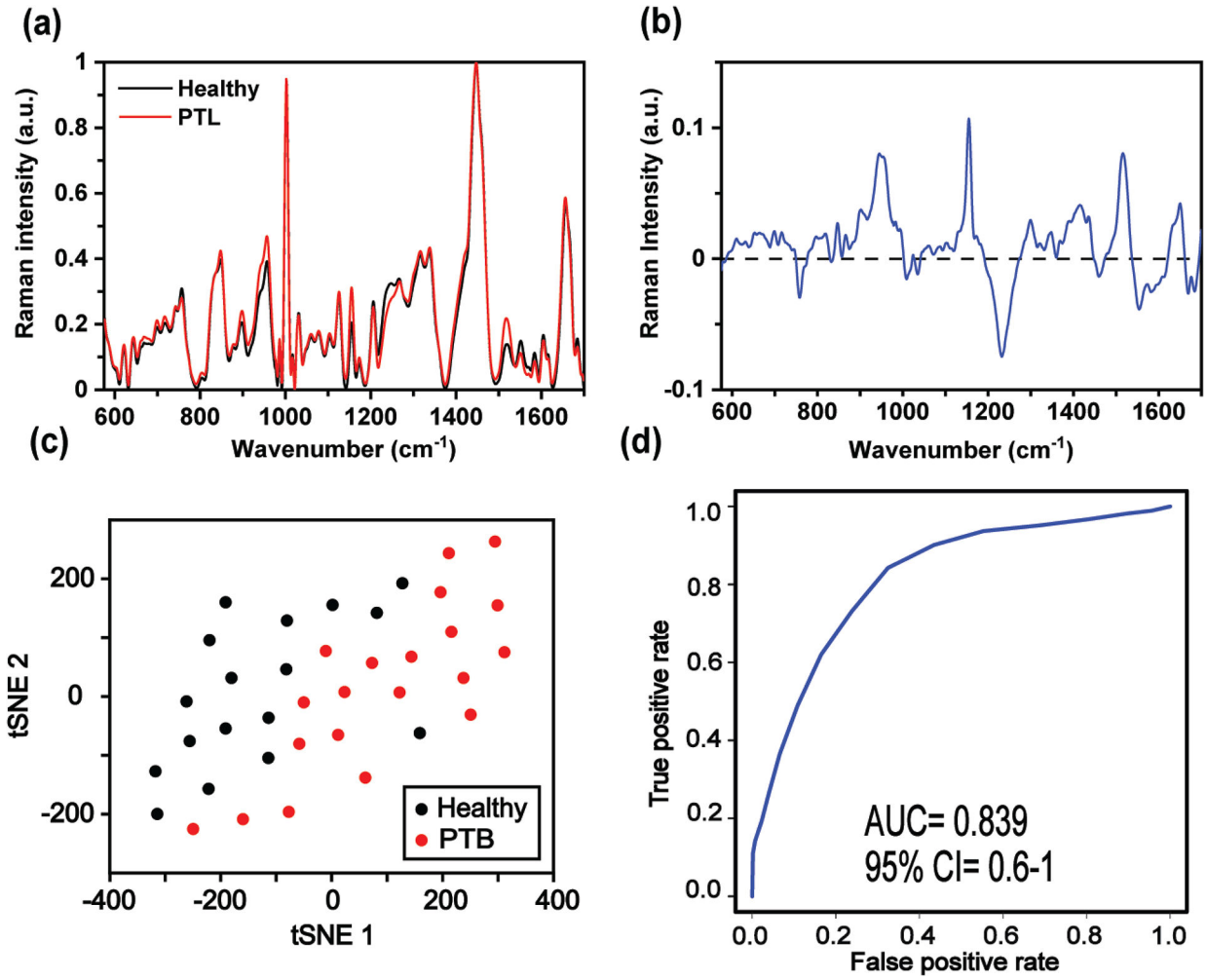


Figure 1.

Raman spectroscopy distinguishes PTB from healthy pregnant patients. (a) Mean, normalized Raman spectra of healthy and PTB patient normalized to the 1447 cm^{-1} lipid peak. (b) The difference spectrum is obtained by subtracting the healthy spectrum from the PTB spectrum in Figure 1a. (c) tSNE of healthy and PTB patients showing separation of the two groups. (d) An area under the curve - receiver operating characteristic (AUC-ROC) curve shows our predictive model can differentiate PTB and healthy patients.

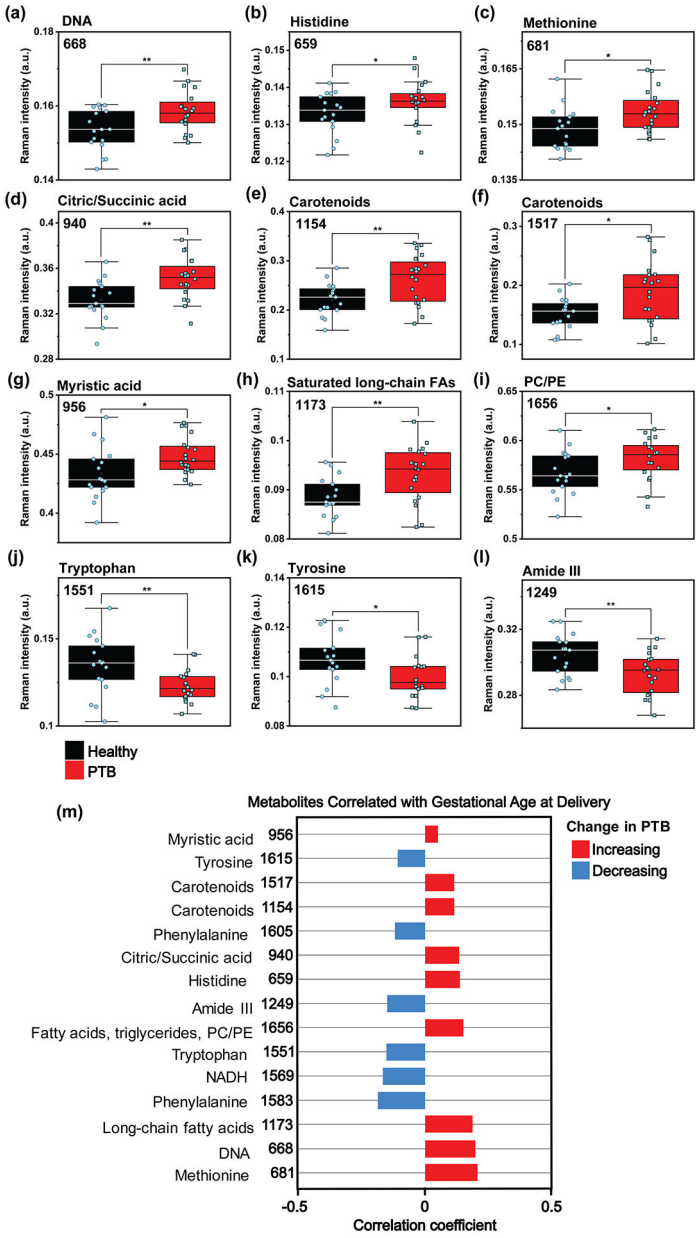


Figure 2.

Key metabolites that distinguish PTB (red) from healthy (black) patients are presented in box plots and correlation coefficient plot to gestational age. (a-i) Metabolites increasing in PTB, including DNA (668 cm^{-1}), histidine (659 cm^{-1}), methionine (681 cm^{-1}), citric/succinic acid (940 cm^{-1}), carotenoids (1154, 1517 cm^{-1}), myristic acid (956 cm^{-1}), long-chain fatty acids (1173 cm^{-1}), fatty acids, triglycerides, and PC/PE (1656 cm^{-1}). (j-l) Metabolites decreasing in PTB, including tryptophan (1551 cm^{-1}), tyrosine (1615 cm^{-1}), and Amide III (1249 cm^{-1}). Here, * represents a $p\ value < 0.05$, and ** represents a $p\ value < 0.01$. (m) Pearson's correlation analysis of significant metabolites to gestational age at delivery. Remaining peaks not included in (a-l) are NADH (1569 cm^{-1}) and phenylalanine (1583, 1605 cm^{-1}).

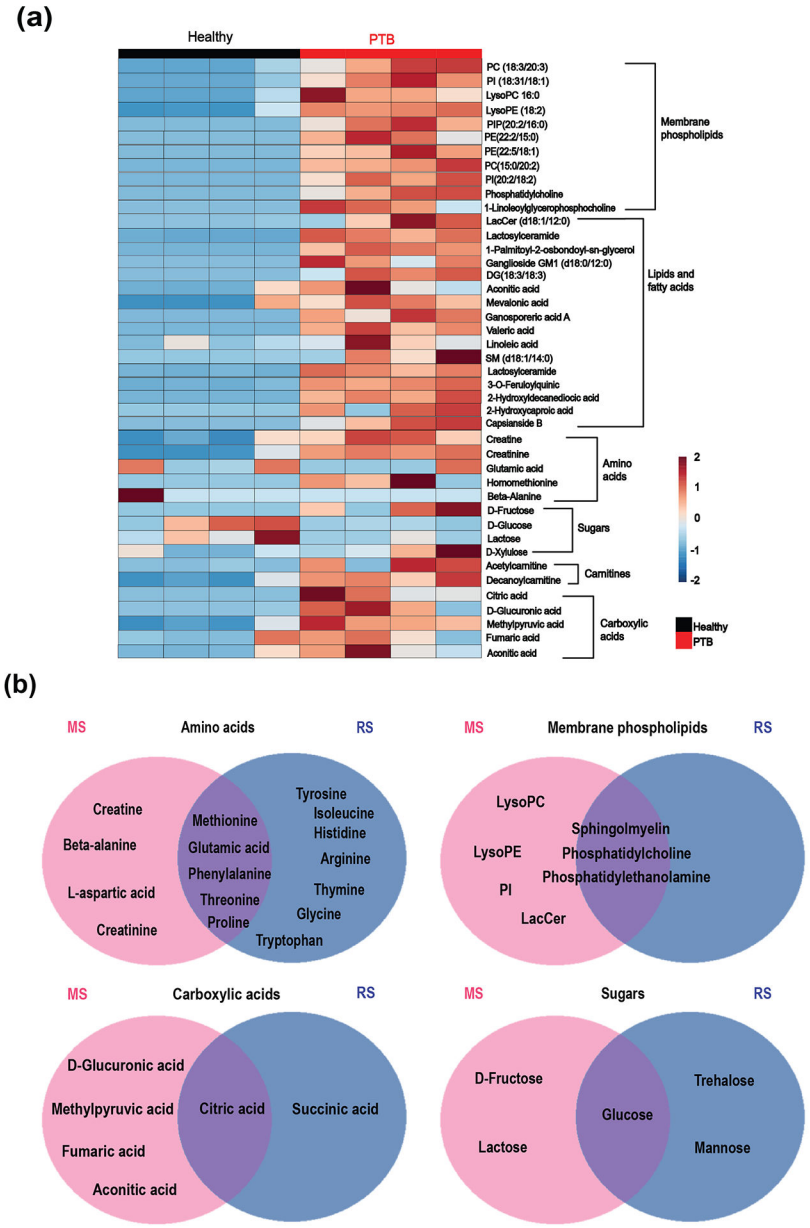


Figure 3.

Mass spectrometry metabolomics distinguishes PTB from healthy. (a) Heat map of statistically relevant metabolites obtained from GC-MS and LC-MS of n = 4 patients per group.(b) Venn diagrams of metabolites in RS and MS. Metabolites in the middle overlap between the two approaches.

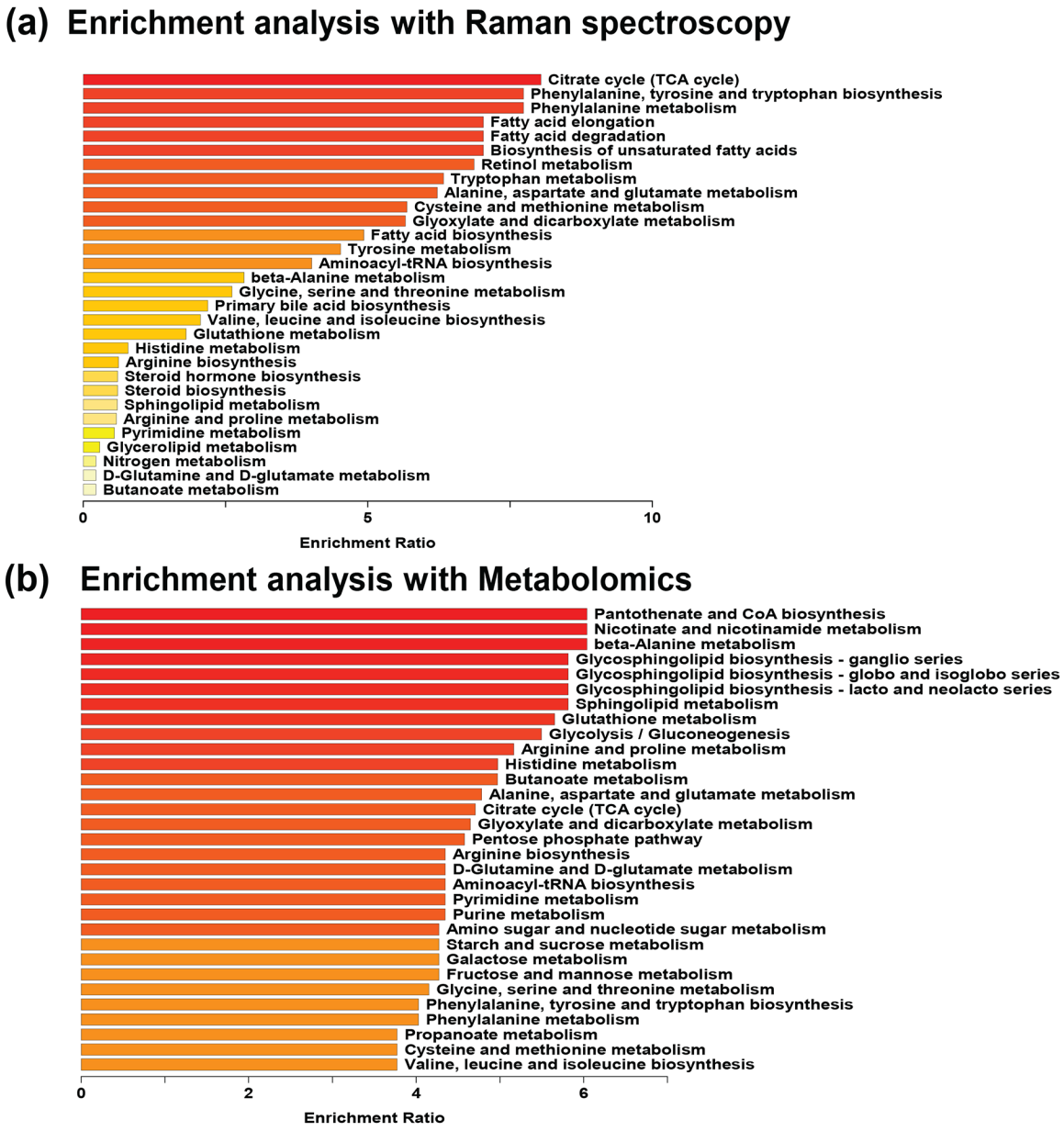


Figure 4.

KEGG pathway enrichment analysis shows metabolic pathways that correspond to the metabolites enriched in the patient samples from (a) Raman measurements and (b) mass spectrometry measurements.

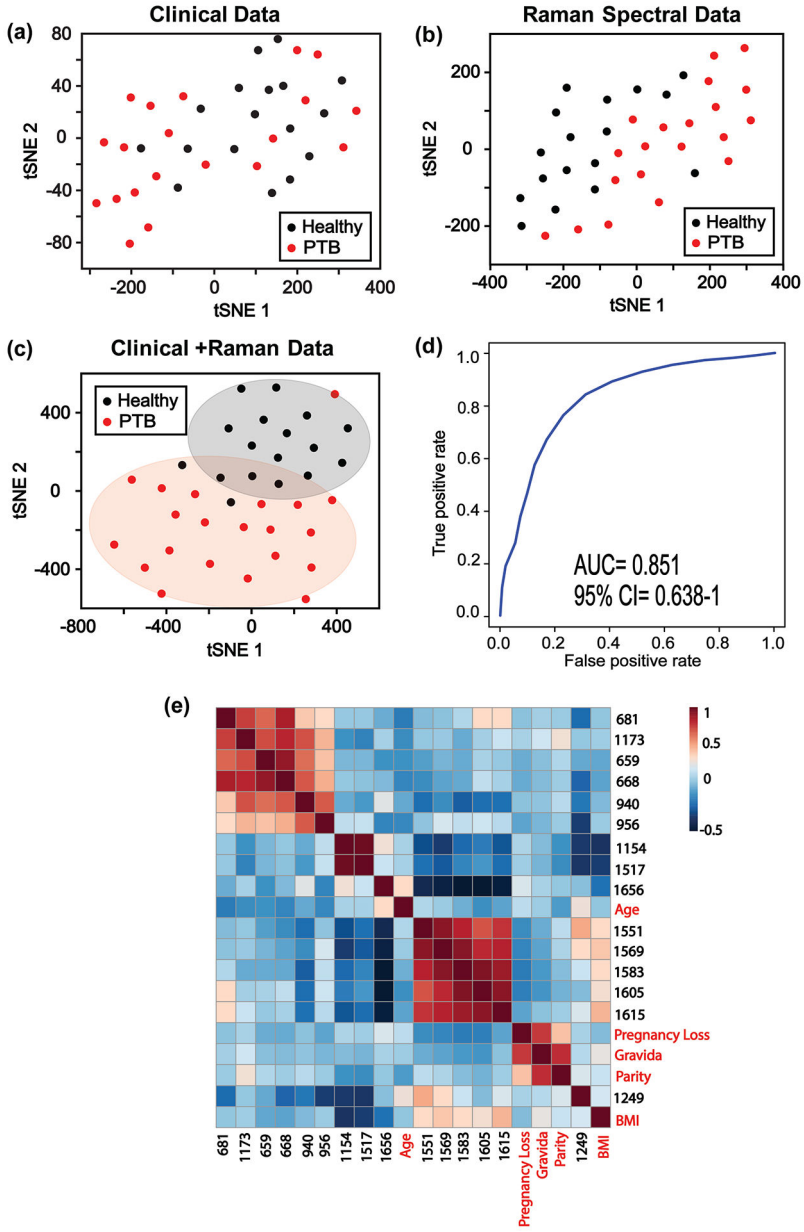


Figure 5.

tSNE of (a) patient clinical data from first trimester healthy and PTB cohort that include maternal age, BMI, gravida, parity, and previous pregnancy losses. (b) Raman data only from Figure 1c, and (c) Raman + clinical information combined present a powerful approach to distinguish patients at high risk of PTB in the first trimester. (d) AUC-ROC analysis shows Raman + clinical data can predict those at risk of PTB with higher accuracy than Raman alone. The AUC was created with a weighted SVM algorithm using the statistically significant peaks from Raman and the five clinical parameters from “a”. (e) Correlation heat map of significant Raman metabolic peaks and clinical data of patients. The Raman metabolic peaks include methionine (681 cm^{-1}), saturated long-chain fatty acids (1173 cm^{-1}), histidine (659 cm^{-1}), DNA (668 cm^{-1}), citric/succinic acid (940 cm^{-1}), myristic acid

(956 cm^{-1}), carotenoids (1154, 1517 cm^{-1}), phospholipids including PC/PE (1656 cm^{-1}), tryptophan (1551 cm^{-1}), NADH (1569 cm^{-1}), phenylalanine (1583, 1605 cm^{-1}), tyrosine (1615 cm^{-1}), and Amide III (1249 cm^{-1}).

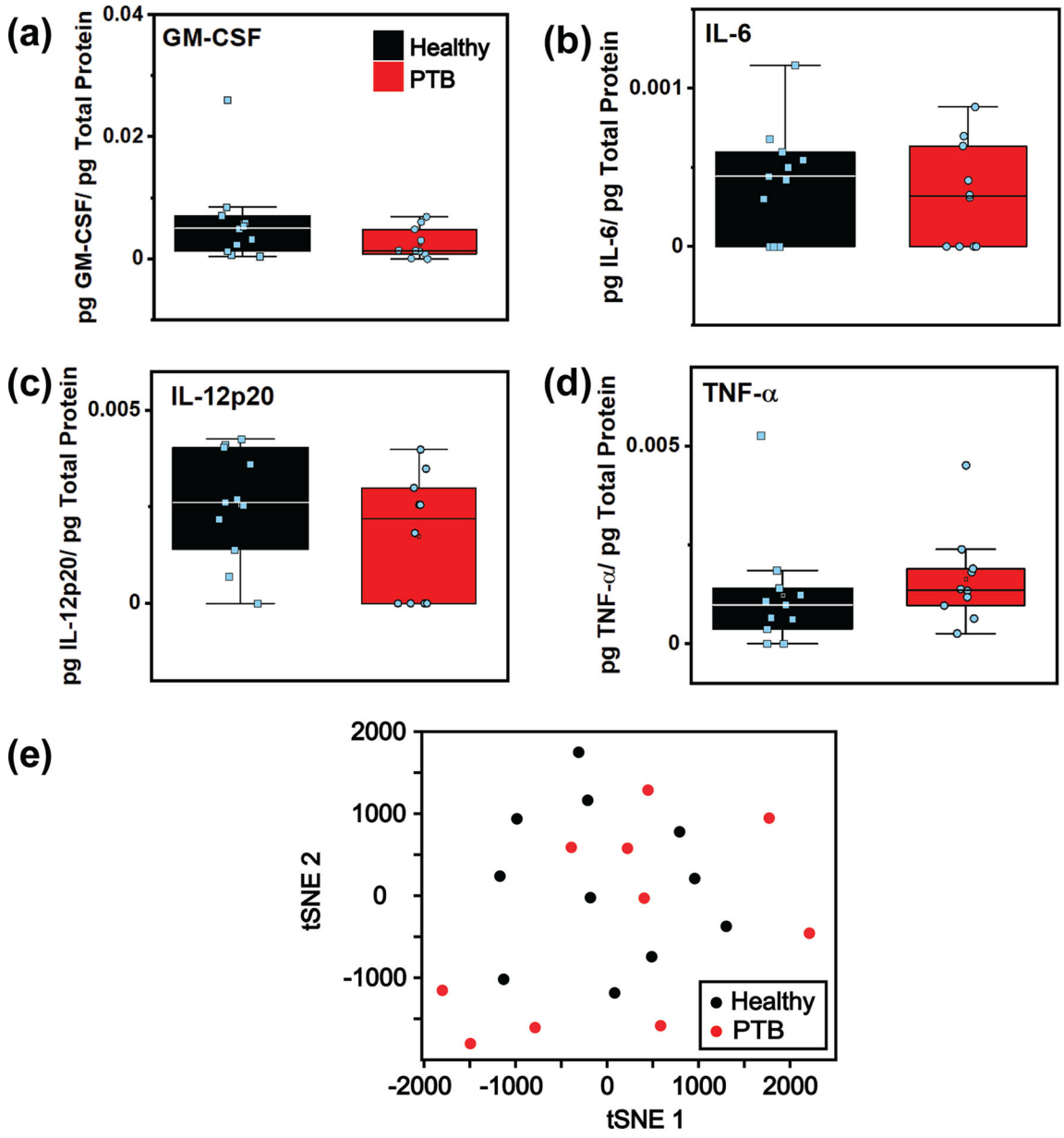
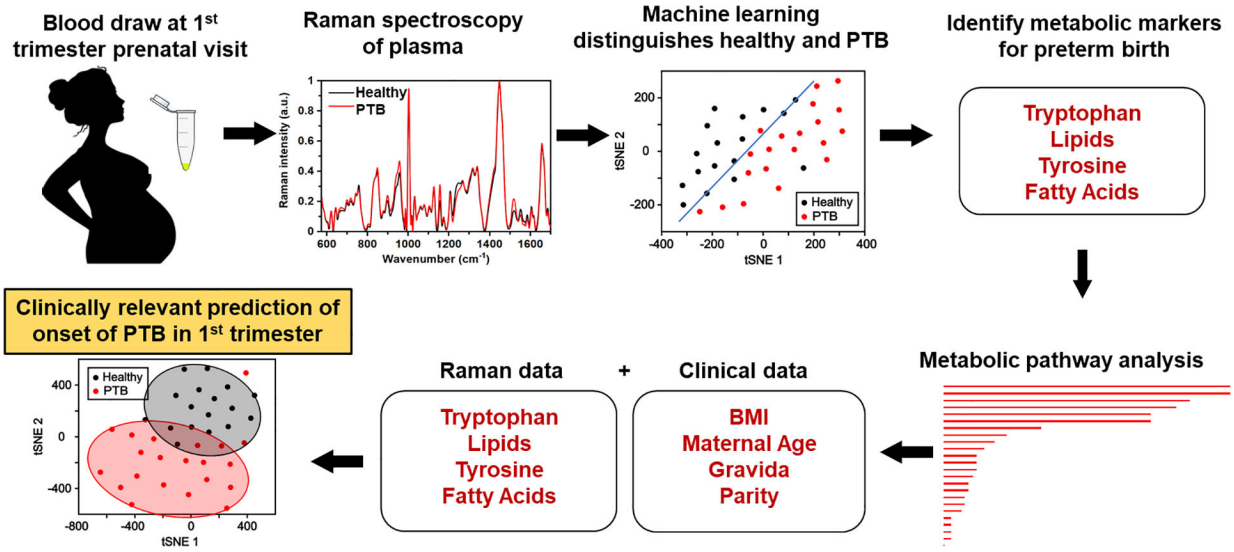


Figure 6.

ELISA of pro-inflammatory cytokines of PTB and healthy patients. (a-d) Box plots of cytokine concentrations of n = 10 PTB and n = 11 healthy patients each, featuring GM-CSF, IL-6, IL-12p20, and TNF-α. Data were normalized to total protein in samples. (e) tSNE of ELISA data shows that cytokine levels in the first trimester are not predictive of the onset of PTB and cannot distinguish the two groups.



Scheme 1.

Schematic representation showing Raman spectra of pregnant patient plasma identify specific metabolites and corresponding metabolic pathways. Raman data combined with clinical data of patients and machine learning enables a predictive model for screening risk of preterm birth in the first trimester.

Table 1.

Demographics and obstetric history of patients. BMI: body mass index.

Clinical Parameter of Patients	Healthy (mean ± SD)	PTB (mean ± SD)	Statistical significance
Number of patients	17	20	-
Maternal age (years)	30.8 ± 5.0	29.7 ± 4.5	0.474890284
BMI (kg/m^2)	29.9 ± 9.9	25.9 ± 7.9	0.192137563
Gestational age at delivery (weeks)	39.1 ± 0.8	34.5 ± 2.2	5.33259E-09
Gravida	2.6 ± 1.4	3.5 ± 2.7	0.298146612
Parity	1.2 ± 0.5	1.4 ± 1.6	0.794293266
Pregnancy loss	0.0 ± 0.0	0.7 ± 1.2	0.026910325
Diabetic (%)	11.7	30	0.189192452
Another comorbidity (%)	5.9	5	0.2481179

Table 2.

Tentative peak assignments for Raman spectral data including the vibrational modes.

Wavenumber (cm^{-1})	Vibrational Modes	Tentative Metabolites	Ref.
621, 1002, 1030, 1583, 1605	C-C twisting, C-C aromatic ring stretching, ring stretching	Phenylalanine	16, 17
643	C-C twisting, C-O-C skeletal mode	Proline	17, 18
659, 1317	C-S stretching, CH_3/CH_2 twisting or bending mode	Histidine	17, 18
668, 805, 1420	C-S stretching of cytosine, C-C stretching, ring breathing mode	DNA	16 18
681	CO^{2-} scissoring and deformation	Methionine	17, 18
699	Choline group, CH_2 rocking	Cholesterol	16, 19, 20
717	C-N symmetric stretch from choline group	Phosphatidylcholine and sphingomyelin	16, 19, 20
743	Backbone vibrations, ring deformation	Thymine	17, 18
756, 1551	Symmetric breathing, C-C stretching, C-C stretching	Tryptophan	17, 18
828, 1615	C-C stretching	Tyrosine	16, 17, 18
848, 1125	C-O-C skeletal stretching	Glucose, glycerol	16 18
878	C-O-C ring, C-O-H bending	Glutamic acid	17, 18
898	CH_2 wagging vibrations from backbone	Glycine	17, 18
940	C-H bending	Citric acid, Succinic acid	18
956	C-H bending	Myristic Acid	19
986	C-N stretching	Arginine	17, 18
1016	C-O-C ring, C-O-H bending	Carbohydrates	16
1060, 1078, 1447	C-C and C-O stretching, CH_2 bending	Lipids	16, 19
1103	$\sigma(CH_2)$ twisting vibrations	Mannose/Trehalose	18
1154, 1517	C-C and C-N stretching, C-C stretching	Carotenoids	16, 20
1173	C-H bending	Saturated long-chain fatty acids	16, 19
1206	NH_3 asymmetric rocking	Amino acid	17, 18
1249	NH_2 group	Amide III	16, 20
1267	C-C stretching, C-H stretching	Fatty acids	16, 19
1300	CH_3/CH_2 twisting or bending mode	Triglycerides	19, 20
1338	CH_3/CH2 twisting/ wagging mode	Threonine	17
1355	CH_3/CH_2 wagging	Isoleucine	17
1569	COO^- group	NADH	16, 18
1656	C-C and C-O stretching	Unsaturated Lipids, Phosphatidylcholine, phosphatidylethanolamine	16, 19, 20
1695	C=C and C=O stretching	Amide I	16