Distribution Agreement

Signature:

______________________________ _______________

Jeffrey M. Collins Date

Plasma Metabolomics in the Diagnosis of Pulmonary Tuberculosis and Early Identification of Multidrug Resistance

By

Jeffrey M. Collins
Master of Science

Clinical Research

____
Henry M. Blumberg, M.D.
Advisor

____
Thomas R. Ziegler, M.D.
Advisor

____
Dean P. Jones, Ph.D.
Committee Member

____
Russell R. Kempker, M.D., M.Sc.
Committee Member

____
Mitchel Klein, Ph.D.
Committee Member

____
Muna Qayed, M.D., M.Sc.
Committee Member

Accepted:

____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

____
Date

Plasma Metabolomics in the Diagnosis of Pulmonary Tuberculosis and Early Identification of Multidrug Resistance

By

Jeffrey M. Collins
B.A., University of Michigan, 2005
M.P.H., Case Western Reserve University, 2011
M.D., Case Western Reserve University, 2011

Advisor: Henry M. Blumberg, M.D.
Advisor: Thomas R. Ziegler, M.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2017

Abstract

Plasma Metabolomics in the Diagnosis of Pulmonary Tuberculosis and Early Identification of Multidrug Resistance

By Jeffrey M. Collins

Introduction: The lack of peripheral blood biomarkers for active tuberculosis (TB) disease is a major obstacle to development of point of care tests for diagnosis and treatment monitoring. Plasma high resolution metabolomics (HRM) is an innovative method to discover and evaluate peripheral blood biomarkers such as those derived from the cell envelope of *Mycobacterium tuberculosis (Mtb)*.

Methods: We compared plasma HRM from 17 patients with active TB disease to 16 household contacts without active TB. We used a suspect screening approach to identify metabolites matching known lipids from the *Mtb* cell envelope based on retention time and accurate mass matches. *Mtb* lipid suspects were tested for association with disease status using the linear model for microarray data. To identify biomarkers of treatment response, we performed plasma HRM on 61 patients at the time of diagnosis and after 4 weeks and 8 weeks of anti-TB therapy. We included 17 patients later found to have multidrug resistant (MDR) TB and 44 patients with drug susceptible TB. A mixed effects model was used to evaluate the effect of MDR status on metabolite intensity and the change in intensity over time.

Results: Plasma HRM identified four *Mtb* lipid suspects significantly increased in active TB patients and undetectable in most household contacts: phosphatidylglycerol (PG), monoacylglycerophosphoinositol (Lyso-PI), monoacylated diacylglycerophosphoinositolmonomannoside (Ac1PIM1), and monoacylglycerophosphoinositolmonomannoside (Lyso-PIM1) ($p<0.001$ for all). These metabolites provided excellent classification accuracy for active TB disease (AUC 0.95). During treatment, the intensity of Lyso-PIM1 was significantly lower in patients with MDR TB compared to those with drug susceptible TB disease ($p=0.01$). Of the other 29 *Mtb* lipid suspects significantly different between MDR TB and drug susceptible TB patients receiving treatment (raw $p\leq0.05$), 27 were also lower in MDR TB patients.

Conclusions: Lyso-PIM1, Lyso-PI, PG and Ac1PIM1 provided excellent classification accuracy for the diagnosis of active TB disease. Additionally, the intensity of Lyso-PIM1 was significantly lower in MDR TB patients during treatment. If confirmed in large studies, further development of these biomarkers could lead to a rapid point of care test for diagnosis and treatment response in active TB disease.

Plasma Metabolomics in the Diagnosis of Pulmonary Tuberculosis and Early Identification of Multidrug Resistance

By

Jeffrey M. Collins
B.A., University of Michigan, 2005
M.P.H., Case Western Reserve University, 2011
M.D., Case Western Reserve University, 2011

Advisor: Henry Blumberg, M.D.
Advisor: Thomas Ziegler, M.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Clinical Research
2017

Table of Contents

List of Tables

List of Figures

Introduction

Tuberculosis (TB) is the leading cause of death due to an infectious disease worldwide; the World Health Organization (WHO) estimated in 2015 there were 10.4 million new TB cases and 1.8 million deaths due to TB (1). Significant challenges in TB control efforts include the difficultly in detecting active TB cases (more than a third of estimated cases are never diagnosed) and lack of biomarkers to assess treatment response (2). There is an urgent need for effective point of care, non-sputum based tests that could be performed at the primary care level to reliably diagnose active TB, monitor response to antibiotic therapy and test for cure. However, the lack of TB biomarkers is a major obstacle to the development of such diagnostic and monitoring tests.

Plasma high-resolution metabolomics (HRM) is an innovative and promising method to discover and evaluate novel peripheral blood biomarkers (3). Walker et al. define the human metabolome as the biologic system that includes "all endogenous metabolites, the chemicals from human–environment interaction, and the reactants arising from the interaction of these compounds with enzymatic and bacterial processes occurring within multiple body components" (4). Plasma HRM utilizes liquid chromatography and high resolution mass spectrometry (LC-HRMS) combined with advanced computational methods in biostatistics and bioinformatics to routinely detect over 10,000 small molecule metabolites in biofluids, including amino acids, lipids, dietary chemicals and the metabolic end products of host-pathogen interactions (4). This technique significantly improves the limit of detection of small molecules and may be crucial for the detection of low-abundance ions such as those derived from the cell envelope of *Mycobacterium tuberculosis (Mtb)* with potential for biomarker development.

The primary purpose of this study was to evaluate the use of plasma HRM to identify *Mtb* cell envelope-derived small molecule metabolites in patients with active pulmonary TB disease and determine whether these biomarkers can discriminate active TB cases from a group of household contacts without TB disease. We further sought to determine whether *Mtb* cell envelope metabolites detected in plasma of patients with active TB disease decline over time among those receiving appropriate anti-TB drug therapy. This pilot study analyzed plasma samples from persons with active pulmonary TB enrolled as part of a randomized controlled trial (RCT) of high dose vitamin D_3 (5). Cross-sectional comparisons for biomarker discovery were made between 17 cases of active pulmonary TB and 16 household contacts. Additionally, plasma HRM over the course of treatment was performed in 44 individuals with drug susceptible pulmonary TB and 17 multidrug resistant (MDR) TB cases. We hypothesized that metabolites derived from the *Mtb* cell envelope would differentiate cases of active pulmonary TB disease from their asymptomatic household contacts and that these metabolites would decline over time in those receiving appropriate anti-TB drug therapy (i.e. in drug susceptible TB cases).

Background

Tuberculosis Diagnostics

The reliable and timely detection of active TB disease remains a major challenge in global TB control efforts. Of the 10.4 million active TB cases estimated to have occurred in 2015 by the WHO, only 6.1 million cases were detected (1). While a variety of factors contribute to low TB case detection rates, the lack of a simple and accurate point of care test to detect active TB disease in a consistent and efficient manner plays a major role (6). Modeling studies indicate improved diagnostic tests have the ability to reduce annual TB mortality by up to 625,000 deaths per year (7).

Early identification of active pulmonary TB cases is a critical step to prevent the ongoing cycle of TB transmission between individuals (8). Transmission of *Mtb* occurs when respiratory droplet nuclei are expelled by a person with active pulmonary TB disease and inhaled by another person (9). As the disease progresses, the burden of *Mtb* organisms in the lungs of the infected host increases, resulting in a greater likelihood of transmission to close contacts (10). Furthermore, failure to diagnose the disease promptly prolongs the infectious interval of active TB cases. Therefore, persons with active pulmonary TB disease must be promptly diagnosed and put in respiratory isolation to minimize transmission (9).

The poor performance characteristics of currently available TB diagnostic tests, especially smear microscopy, which is frequently used in low and middle income countries, is a major contributor to poor case identification and prolonged diagnostic delays. Diagnostic sputum culture is the gold standard for diagnosis of active pulmonary TB disease, but can take 2-6 weeks for definitive results, is expensive and is uncommonly available in resource-limited settings. This results in missed diagnoses, ongoing transmission of *Mtb* from infectious patients and unnecessary isolation and empiric treatment of some patients later found not to have TB (i.e. culture negative). The most commonly used test for rapid diagnosis of active pulmonary TB disease remains direct sputum smear microscopy with Ziehl-Neelsen staining for presence of acid-fast bacilli (AFB) (11). However, the sensitivity for detection of active pulmonary TB is poor: 50% in immunocompetent patients (12) and less than 30% among immunocompromised patients with HIV co-infection (13). Sensitivity of smear microscopy is even lower for the diagnosis of extra-pulmonary TB, which accounted for 15% of global TB cases in 2015 (1). A rapid molecular diagnostic test for TB, the Xpert® MTB/RIF assay, is able to detect up to 70% of AFB sputum smear negative cases and 98% of smear positive cases (6). However, it is typically only available at referral health centers globally and remains unaffordable in many resource-limited settings (14). Furthermore, all tests require analysis of sputum specimens, which many patients (particularly young children) are unable to produce.

Biomarkers of Active TB Disease

There is an urgent and pressing need for a rapid, non-sputum based test that can identify patients with active TB disease and rule out the disease in patients without active TB disease (6, 15). A peripheral blood-based biomarker signature for active TB would greatly facilitate the development of a point of care test. Diagnostic assays performed on peripheral blood samples have thus far had limited utility in the diagnosis of active TB. Of the hundreds of studies identifying potential biomarkers of active TB disease, only a small fraction have been validated in independent, prospective studies (16). Transcriptomic signatures have shown an ability to discriminate active TB cases from both healthy controls, as well as clinically ill patients suspected of having active TB infection, but ultimately diagnosed with something else (17). However, the groups of genes involved in transcriptional signatures frequently vary in different populations (17). Several tests have been developed to measure *Mtb*-specific antibodies and all demonstrate poor sensitivity and specificity (18). Interferon-gamma release assays do not directly measure *Mtb* antigens and are unable to differentiate latent TB infection (LTBI) from active disease (19). To date, the urine-based lateral flow assay for the *Mtb*-specific cell wall glycolipid lipoarabinomannan is the only biomarker to advance to prospective clinical trials (16). However, these studies demonstrated very low sensitivity except in severely ill patients with advanced HIV (20).

Early Identification of MDR TB

The emergence and spread of highly drug resistant TB is another important challenge to global TB control. An increasing number of active TB cases globally are due to MDR TB, with an estimated 480,000 cases diagnosed worldwide in 2015 (1). MDR TB is defined by resistance to two key first-line antibiotics: rifampin and isoniazid. Patients with MDR TB have poorer outcomes and much higher mortality rates compared to patients with drug susceptible TB (1). Traditionally, response to pulmonary TB treatment is measured by serial analysis of sputum samples (smear and culture) (21). However, 8-week clearance of sputum cultures is an imperfect predictor of long term treatment outcomes (22). Moreover, culture with drug susceptibility testing can take several weeks, meaning patients may be on ineffective therapy for months before drug resistance is discovered. This can lead to additional or amplified drug resistance and increased mortality due to treatment failure. The molecular diagnostic Xpert® MTB/RIF assay can detect baseline resistance to the drug rifampin, which is often a marker of MDR TB, but does not offer information about susceptibility to other TB drugs. Furthermore, genetic material from killed *Mtb* organisms may remain detectable in sputum samples of patients responding to anti-TB therapy, limiting its usefulness as a marker of treatment response. Thus, development of biomarkers indicative of treatment response and drug resistance would greatly enhance the ability to monitor patient response to treatment and could enhance patient outcomes.

Plasma Metabolomics and Biomarker Discovery

Plasma HRM has the potential to detect novel biomarkers of active TB disease, which may be able to reliably diagnose active TB from blood samples (23, 24). Plasma HRM utilizes LC-HRMS to detect at least 10,000 metabolites (>100,000 ions) in biologic samples (23-25). This capacity is obtained by use of accurate mass instruments, analysis of plasma samples in triplicate with rigorous standard operating procedures, and advanced data extraction methods (23-25).

Recent studies utilizing metabolomics to identify potential biomarkers in active TB disease have relied on an untargeted metabolomics approach using human (endogenous) metabolite databases for chemical annotation (26, 27). However, existing chemical libraries allow for annotation of only a small fraction (approximately 2%) of spectra produced by a mass spectrometer in an untargeted metabolomics experiment (28). A focus on human metabolic changes in patients with active TB disease may reveal non-specific alterations in metabolism and tend to overlook metabolites derived from *Mtb* itself. Those studies examining metabolites specific to *Mtb* have employed a highly targeted approach; measuring only a few *Mtb*-specific metabolites in biologic samples and potentially limiting the sensitivity of disease detection (29, 30). To our knowledge, only a single study has used metabolomics to measure patient response to antibiotic therapy (31). Similar to studies examining the use of metabolomics for the diagnosis of active TB, this study examined changes in human metabolites present in urine samples (31).

The high sensitivity of HRM allows for simultaneous detection and semi-quantitative measurement of multiple low-abundance metabolites, such as those produced by degradation of *Mtb* by the infected host (24). Preliminary studies suggest that metabolites found in the *Mtb* cell envelope are detectable in plasma of patients with active pulmonary TB disease and largely absent from their asymptomatic household contacts (23). However, no studies to date have specifically targeted metabolites derived from the *Mtb* cell envelope by comprehensively screening plasma of patients with active TB disease in relation to control subjects without active TB disease. Multiple groups have recently created databases of lipid species derived from the *Mtb* cell envelope, characterized by retention time and mass to charge ratio (m/z), to identify potential biomarkers of active TB disease (32, 33). If a series of these *Mtb*-derived metabolites were found in patients with active TB disease, they would have great potential for development as biomarkers that could be used to create a diagnostic test for rapid detection of active TB disease from peripheral blood samples.

Methods

The first aim of this study was to evaluate whether metabolites derived from the *Mtb* cell envelope could be detected in the plasma of patients with active pulmonary TB disease and to determine whether detected metabolites can differentiate TB patients from a group of asymptomatic household contacts. This study aim was hypothesis generating, as no specific *Mtb* metabolites were targeted for analysis.

The second study aim was to assess whether *Mtb* metabolites found in plasma of patients with active TB disease can identify patients with MDR TB when followed over time in patients receiving anti-TB drug therapy. We hypothesized that *Mtb* metabolites differentiating active TB cases from household contacts would decrease over time in patients with drug sensitive TB (receiving effective antibiotic therapy) and would remain elevated in patients with MDR TB (later found to be receiving ineffective antibiotic therapy). We also explored whether changes in the concentration of other *Mtb* metabolites during treatment differentiated patients with MDR TB from those with drug sensitive TB. This aspect of the analysis was also hypothesis generating.

Study Design

A cross-sectional study design was used to identify potential *Mtb* cell envelope-derived biomarkers in plasma of active pulmonary TB patients. Metabolite concentrations in this group near the time of diagnosis were compared to a group of asymptomatic household contacts attending the TB clinic with the active TB patient. In measuring response to treatment, a retrospective cohort design was used to assess the change in plasma metabolite concentrations over the course of TB treatment in patients with drug susceptible TB compared to those with MDR TB.

Study Population

Patients with Active Pulmonary TB Disease

For the biomarker discovery analysis, we examined plasma HRM for 17 patients with active TB disease selected from a double blind, randomized, controlled, prospective trial of high-dose cholecalciferol treatment of patients with pulmonary TB conducted in the country of Georgia (clinicaltrials.gov identifier NCT00918086) (5). Inclusion criteria for patients included age ≥ 18 years, newly diagnosed TB as determined by a positive AFB sputum smear, and later confirmed by a positive sputum culture for *Mtb* (performed at the Georgian National TB Reference Laboratory (NRL)) within 7 days of treatment with anti-TB drug therapy. Patients receiving more than 7 days of anti-TB therapy, as well as those with hypercalcemia, nephrolithiasis, hyperparathyroidism, sarcoidosis, history of organ transplant, liver cirrhosis, end stage renal disease, serum creatinine greater than 0.250 mmol/L, cancer in past 5 years, seizures, current pregnancy or lactation, current use of cytotoxic or immunosuppressive drugs, or current incarceration were excluded.

For the repeated measures analysis, plasma HRM was performed on 61 patients from the same study cohort at the time of diagnosis and after 4 weeks and 8 weeks of anti-TB therapy (Figure 1). All patients were treated with the WHO standard first line antibiotic treatment regimen, which includes rifampin, isoniazid, pyrazinamide, and ethambutol (21). All enrolled study participants underwent conventional sputum culture, and all *Mtb* isolates were sent for drug susceptibility testing (DST). In our study, baseline and serial plasma was obtained before DST results diagnosing MDR TB were available (> 8 weeks after initiation of conventional anti-TB therapy); thus, the subjects with MDR TB were receiving ineffective drug therapy during plasma sampling for HRM reported here. We included 17 patients later found to have MDR TB and 44 patients with confirmed drug susceptible TB. MDR TB was defined as infection with an *Mtb* isolate resistant to both isoniazid and rifampin.

Household Contacts

Sixteen household contacts were enrolled from persons accompanying patients to the TB treatment facility at the time of study enrollment. The group was primarily made up of close relatives of enrolled patients. To be eligible for enrollment, participants were required to have a negative screen for symptoms suggestive of TB disease or any other acute illness and documented negative AFB sputum smear and culture.

Sputum Culture and Drug Susceptibility Testing

Serial sputum specimens at weeks 2, 4, 6, 8, 12 and 16 after RCT entry were obtained from each patient confirmed to have active pulmonary TB disease. All sputum cultures were performed at the Georgian NRL using standard methodologies as previously described (34). Positive cultures were confirmed to be *Mtb* complex using phenotypic tests and the MTBDR*plus* assay, as outlined (35). Drug susceptibility testing was done using absolute concentration method on solid media with standard methodology as previously described (34).

Plasma Sample Collection

Peripheral blood samples were obtained by venipuncture from all subjects with TB disease as well as an asymptomatic household contact without TB disease at the time of diagnosis. Those patients with active TB disease had blood samples taken after 4 weeks and 8 weeks of treatment. Blood was collected in ethylenediaminetetraacetic acid (EDTA)-containing tubes, centrifuged and isolated plasma immediately stored frozen at -80º C. Samples were subsequently shipped on dry ice from the NRL in Tbilisi, Georgia to Emory University, Atlanta, GA, USA. Samples remained frozen during transit, and were kept at -80º C in Atlanta prior to metabolomics analysis.

Plasma Metabolomics Analysis

Thawed plasma (65 mL) was treated with 130 ml acetonitrile (2:1, v/v) containing an internal isotopic standard mixture (3.5 mL/sample), as previously described (23, 36). The internal standard mix for quality control consisted of 14 stable isotopic chemicals covering a broad range of chemical properties represented in small molecules (36). Samples were mixed and placed in ice for 30 min prior to centrifugation for 10 min (16,100 x g at 4º C) to remove protein. The supernatants (10 mL), for each LC-HRMS analysis were then loaded onto an autosampler maintained at 4º C and analyzed in triplicate using a LTQ-Velos Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA) and C18 chromatography (Higgins Analytical, Targa, Mountain View, CA, USA, 2.1 x 10 cm). Elution was obtained with a formic acid/acetonitrile gradient at a flow rate of 0.35 ml/min for the initial 6 min and 0.5 ml/min for the remaining 4 min. The first 2-min period consisted of 5% solution A [2% (v/v) formic acid in water], 60% water, 35% acetonitrile, followed by a 4-min linear gradient to 5% solution A, 0% water, 95% acetonitrile. The final 4-min period was maintained at 5% solution A, 95% acetonitrile. The mass spectrometer was set to collect data from m/z 85 to 2000 Daltons over a 10-minute chromatography period. Electrospray ionization was used in positive ion mode for detection, as outlined (23, 25). Leading software packages apLCMS and xMSanalyzer were used for data extraction as previously outlined to provide feature tables containing accurate m/z values, retention times and intensities, as previously outlined (37, 38). Metabolite intensity is a semi-quantitative measure of relative abundance of specific metabolites. All metabolite intensity values were log2 transformed for comparison between groups.

Metabolite Identification

A suspect screening approach was used to identify potential *Mtb* lipid metabolites present in plasma samples (39). Metabolites identified from participant plasma samples were characterized by their retention time on the liquid chromatography column, as well as m/z ratio. Because *Mtb* cell envelope lipids are expected to have a longer retention time, those molecular features with retention time < 3 minutes were excluded to limit false matches by m/z ratio. A previously published *Mtb* lipid library, "*Mtb* LipidDB" was used to annotate metabolites as *Mtb* lipid suspects based on accurate mass matches (32). The database search included M+H, M+Na and M-H_2O+H adducts for each lipid species with a mass accuracy threshold of 10 parts-per-million (ppm).

Data Analysis

All analyses were performed using R version 3.3.0. Comparisons of descriptive statistics and clinical characteristics between pulmonary TB cases and household contacts and MDR TB and drug susceptible TB cases respectively were performed using a two-tailed independent t-test for continuous variables and a Fisher exact test for categorical variables.

Mtb Lipid Biomarker Discovery

The log2 intensities of *Mtb* lipid suspects were compared between active pulmonary TB cases and household contacts using the linear model for microarray data (LIMMA) (40). Two-way hierarchical cluster analysis (HCA) was performed using all metabolites with a raw p-value ≤ 0.05 to determine discrimination between pulmonary TB cases and household contacts. The Benjamini-Hochberg false discovery correction of 20% was then applied to correct for multiple comparisons and to identify those metabolites of greatest interest (41). Logistic regression analysis using the log2 intensity value of the most significant metabolites was then used to construct an ROC curve to determine the classification accuracy. All analyses were performed using xMSpanda; an in-house R package (Uppal K et al. *submitted, in review*).

Biomarkers for Identification of MDR TB

The log2 intensity of *Mtb* lipid metabolites was compared between persons with MDR TB and persons with drug susceptible TB over time using a mixed model with a subject-specific random effect. For each *Mtb* lipid suspect we modeled the effect of MDR status, time point of analysis and the interaction between these variables.

Model of effect of MDR status: $\text{Log}_2\,(\text{Metabolite Intensity})_{ij} = \beta_0 + \beta_1\,(\text{MDR Status}) + \beta_2\,(\text{Week 4}) + \beta_3\,(\text{Week 8}) + \mathbf{b}_{0i} + \varepsilon_{ij}$

Model of effect of MDR status over time: $\text{Log}_2\,(\text{Metabolite Intensity})_{ij} = \beta_0 + \beta_1\,(\text{MDR Status}) + \beta_2\,(\text{Week 4}) + \beta_3\,(\text{Week 8}) + \beta_4\,(\text{Week 4}) * (\text{MDR Status}) + \beta_5\,(\text{Week 8}) * (\text{MDR Status}) + \mathbf{b}_{0i} + \varepsilon_{ij}$

$\text{Log}_2\,(\text{Metabolite Intensity})_{ij}$ represents the measure of metabolite intensity for the j^{th} observation of the i^{th} subject; $\mathbf{b}_{0i}$ is the subject-specific random effect, and ε_{ij} is within-subject residual error. The effect of MDR status and the interaction of MDR status and time point were evaluated using a likelihood ratio test (42). For metabolites identified in the cross-sectional discovery analysis as differentiating pulmonary TB cases from household contacts, a raw p-value of 0.05 was used to determine statistical significance of model coefficients. For discovery analysis of *Mtb* lipids not previously specified, p-values were adjusted to a FDR of 20% as previously described (41). All analyses were performed using the lme4 package in R (42).

IRB Approval

This study was approved by the Emory University and Georgian National Center for Tuberculosis and Lung Disease Institutional Review Boards.

Results

Plasma *Mtb* Lipids Distinguish Active TB Cases from Household Contacts

Participant Characteristics

As shown in Table 1, the mean age (33.4 vs 38.6, p=0.18) and male sex (59% vs 38%, p=0.22) did not differ significantly in the 17 active TB cases compared to the 16 household contacts in the cross-sectional analysis. There were significantly more smokers in the active TB group compared to the household contacts (76% vs 38%, p=0.02; Table 1). All enrolled TB patients had pulmonary disease and none had any extrapulmonary manifestations. Three (18%) of the TB cases had MDR TB and only two (12%) had evidence of cavitation on x-ray.

High Resolution Metabolomics

Following data extraction with apLCMS and xMSanalyzer, 33,262 unique m/z features (metabolites) were detected. Filtering for median coefficient of variation ≤ 100% yielded 32,975 remaining molecular features. Testing these metabolites for accurate mass matches to the 1,696 unique lipid species contained in the *Mtb* lipid database "*Mtb* LipidDB" (32) yielded 2,470 mass matches within 10 ppm. Metabolites exhibiting retention time ≤ 3 minutes with ≥ 50% missing values in the active TB group were considered unlikely to represent *Mtb* lipid metabolites and were removed from the analysis, leaving 867 *Mtb* lipid suspects.

Mtb Lipids Associated with Active TB Disease

Using LIMMA to test the association of log2 intensity values of *Mtb* lipid matches with active TB disease produced 69 suspects with raw p ≤ 0.05 (Figure 2). Application of a false discovery rate correction of ≤ 20% identified four *Mtb* lipid suspects significantly increased in active TB patients and undetectable in most household contacts (raw p<0.001

for all) (Figure 3). The four most significant *Mtb* lipids matched the m/z and retention time of phosphatidylglycerol (PG), monoacylglycerophosphoinositol (Lyso-PI), monoacylated diacylglycerophosphoinositolmonomannoside (Ac1PIM1), and monoacylglycerophosphoinositolmonomannoside (Lyso-PIM1). All four metabolites were significantly increased in patients with pulmonary TB compared to household contacts, consistent with potential biomarkers of active TB disease. At least one of the four most significant *Mtb* metabolites was present in all persons with active TB disease and two or more were detectable in 14 (82%) of 17 TB cases. Among the household contacts, eight (50%) had one detectable *Mtb* metabolite and only one (6%) had two detectable metabolites.

Of the 69 *Mtb* lipid suspects significantly different between active TB cases and household contacts at raw $p \leq 0.05$, 38 were increased in patients with active TB. Most belonged to the lipid classes Ac1PIM1, monoacylated diacylglycerophosphoinositoldimanosides (Ac1PIM2), Diacylglycerophosphoglycerophosphodiradylglycerols (CL) and Diacylglycerophosphoinositoldimannosides (PIM2). Those *Mtb* lipid suspects significantly increased in patients with active pulmonary TB disease (compared to household contacts) and correlated with at least two of the four most significant metabolites (Spearman correlation coefficient ≥ 0.3) are summarized in Table 2.

Analysis of TB Disease Classification Accuracy

Two-way HCA of the 69 *Mtb* lipid metabolites significantly different between TB cases and household contacts at raw $p \leq 0.05$ revealed near complete separation of pulmonary TB cases and household contacts without TB (Figure 4). Using a logistic regression model, the log2 intensity value of the four most significant metabolites (Ac1PIM1, Lyso-PIM1, PG and Lyso-PI) also provided excellent classification accuracy of active TB cases (Figure 5). The receiver operating characteristic curves demonstrates an area under the curve (AUC) of 0.95 when using all four metabolites and 0.82 when using only the top two metabolites. This performance is similar to the unsupervised HCA analysis using all 69 significant metabolites.

Plasma *Mtb* Lipids and Detection of MDR TB

Participant Characteristics

To examine whether changes in the four identified *Mtb* lipid biomarkers over time were predictive of MDR TB, we examined plasma metabolomics in 44 patients with drug susceptible TB and 17 patients with MDR TB at baseline and serially at week 4 and 8 after study entry. The clinical characteristics between the two groups of active TB disease subjects were similar (Table 3). Patients in the drug susceptible group achieved sputum culture clearance more rapidly by week 8 than patients with MDR TB (92.9% vs 47.1%, $p<0.001$); the difference at week 4 was not statistically significant (54.8% vs 35.3%, $p=0.13$).

High Resolution Metabolomics

Following data extraction with apLCMS and xMSanalyzer, 13,451 unique m/z features (metabolites) were detected by LC-HRMS in the serial samples. Metabolites exhibiting retention time ≤ 3 minutes were removed from the analysis as these are unlikely to represent *Mtb* lipids. Testing the remaining m/z values for accurate mass matches to the 1,696 unique lipid species contained in the *Mtb* lipid database "*Mtb* LipidDB" (32) yielded 464 mass matches within 10 ppm.

Changes in Ac1PIM1, Lyso-PIM1, PG and Lyso-PI Over Time

The four *Mtb* cell envelope metabolites identified in plasma by the cross-sectional discovery analysis as markers of active TB disease were included in the repeated measures analysis. All four metabolites remained detectable in the plasma obtained from patients with both drug susceptible TB and MDR TB after eight weeks of treatment (Figure 6). The intensity of the *Mtb* lipid Lyso-PIM1 was significantly lower in patients with MDR TB compared to those with drug susceptible TB disease ($p=0.01$). This effect was stable at each time point and there was no significant change over time ($p=0.77$ for interaction between time point and MDR status). Ac1PIM1, PG and Lyso-PI did not demonstrate any statistically significant differences in intensity level between patients with drug susceptible TB and those with MDR TB, and metabolite intensities did not change over time based on MDR status.

Mtb Cell Envelope Lipids and Detection of MDR TB

Using a mixed effects model, there were 30 *Mtb* lipid suspects in plasma with significantly different intensity levels between patients with drug susceptible TB and those with MDR TB at raw p-value ≤ 0.05. Of these 30 metabolites, 28 had a significantly greater intensity in plasma from drug susceptible TB patients compared to MDR TB patients. Analogous to Lyso-PIM1, this was not a time-dependent effect and the metabolite intensity increase was similar at all time points. The lipid classes significantly increased in plasma from drug susceptible TB patients compared to those with MDR TB included Ac1PIM2, Ac1PIM3, Ac2PIM2, PIM5, trehalose monomycolate, glucose monomycolate and diacylated sulfolipids (Table 4). When using a false discovery rate correction of 20%, none of these metabolites was statistically significant when comparing drug susceptible TB patients to MDR TB patients. There were no *Mtb* lipid suspects that demonstrated consistent and statistically significant changes in metabolite intensity over time based on MDR TB status.

Discussion

The results of this pilot study demonstrate that by using HRM numerous *Mtb*-derived cell envelope metabolites are detectable in the plasma of patients with active pulmonary TB disease. Thus, these can be further explored as potentially useful biomarkers to differentiate TB cases from asymptomatic household contacts in subsequent larger trials. This study is the first, to our knowledge, to use plasma HRM to identify a metabolomics signature targeting the *Mtb* cell envelope in patients with active TB disease in relation to controls without evident active TB from the same household. We found that the *Mtb* cell envelope-derived metabolites Lyso-PIM1, Lyso-PI, PG and Ac1PIM1 were significantly increased in patients with active pulmonary TB disease and largely absent from their asymptomatic household contacts. These metabolites demonstrated excellent classification accuracy for active TB cases (AUC 0.95) and therefore may have promise for further development as a metabolomics signature of active TB disease.

It is also possible that additional *Mtb* cell envelope metabolites other than Lyso-PI, PG, Ac1PIM1 and Lyso-PIM1 will be important components of a metabolic signature of active pulmonary TB disease. This study was hypothesis generating, and although these top four metabolites produced excellent separation between TB cases and household contacts, it is unlikely this performance would be matched in an independent cohort of active TB patients. It will therefore be essential to test these four metabolites as a diagnostic tool in a larger, independent cohort of patients with active pulmonary TB disease. A confirmatory study of this nature would better capture the true diagnostic accuracy of these four metabolites. If the test characteristics are not confirmed in an independent population, then it may be important to include other *Mtb* metabolites in the metabolic signature, such as those not reaching significance using the FDR correction of 20%.

Current diagnostic tests for active TB disease require patients to produce a sputum specimen and generally require laboratory facilities not available in many resource-limited areas. Novel biomarkers that are reliably and specifically detectable in patients with active TB disease are needed to develop new diagnostic tests capable of rapidly detecting active TB from peripheral blood samples. Peripheral blood biomarkers from the *Mtb* cell envelope such as those identified in our study have potential for development into a point of care diagnostic test that can be used in resource-limited areas where laboratory facilities are often absent.

Previous studies using untargeted metabolic profiling to identify peripheral blood biomarkers of active TB disease have primarily focused on identifying differences in human endogenous metabolites associated with active TB disease (23, 26, 27). The metabolite databases typically used to annotate metabolites, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG), the Human Metabolite Database (HMDB) and METLIN, contain few metabolites derived from infecting microorganisms (43-45). Although a pilot HRM study by our group identified an *Mtb*-derived cell envelope metabolite, trehalose-6-mycolate, using METLIN (23), metabolites from the *Mtb* organism are not comprehensively cataloged by these open source databases. Thus, for this targeted study, we used LC-HRMS and data from a previously published *Mtb* lipid library to annotate metabolites as *Mtb* lipid suspects based on accurate mass matches (32). Additionally, the HRM method used in this study has the capacity to detect thousands of low-abundance metabolites in plasma; much greater sensitivity than low-resolution metabolomics methods used for analysis of blood, sputum, urine and breath in previous studies (46).

The *Mtb* cell envelope has several lipid species not present in other bacterial or human cells, making these ideal targets for biomarkers of active TB disease (30, 32, 33, 47).

Although two of the metabolites identified in this study can be found in human cells (Lyso-PI and PG) and two in other species of mycobacteria (Ac1PIM1 and Lyso-PIM1), these metabolites are abundant in the cell envelope of *Mtb* (32) and therefore may be more specific to active TB disease. The *Mtb*-derived glycolipid lipoarabinomannan has excellent specificity for active TB disease, but sensitivity is poor (20). A signature of several glycolipids may offer improved sensitivity by targeting several lipid species from the *Mtb* cell envelope. Alterations in human metabolism associated with active TB disease may result from a variety of host response mechanisms and therefore may be present in other pro-inflammatory disease states.

Although only one of the household contacts in this study had detectable concentrations of two of the four most significant *Mtb* lipid metabolites, approximately half of these individuals had detectable levels of at least one of these *Mtb* metabolite matches in plasma. There are several possible explanations for this finding. Given most household contacts would have been heavily exposed to the active pulmonary TB cases enrolled in the study, it is likely that a large proportion of these subjects had LTBI, which was not tested for in this cohort of household contacts. There is increasing evidence that patients with LTBI have ongoing organism replication requiring continuing control by the immune system (48), potentially resulting in detectable *Mtb* metabolites in plasma of these patients as well. Alternatively, detection of these metabolites in individuals without active TB may reflect a lack of specificity to active TB disease. For example, PG and Lyso-PI lipids are also found in human cells, though they are abundant in the cell envelope of *Mtb* (49). In future studies, it will be important to use better defined control groups (e.g. with rigorous negative testing for LTBI) to further refine the series of *Mtb* metabolites diagnostic of active TB disease. These metabolites, for example, could be compared to both patients with confirmed LTBI as well as people without latent or active TB over time to determine whether they are specifically present in active TB disease.

This study also demonstrates that *Mtb* lipid metabolites continue to circulate in the bloodstream of active TB patients for weeks despite appropriate antibiotic therapy. Lyso-PIM1, Lyso-PI, PG and Ac1PIM1 remained detectable in the plasma of patients with both MDR TB and drug susceptible TB, despite receiving adequate first line anti-TB therapy. This suggests that once the metabolites are present, they are not rapidly cleared from the body. Until the current study, investigations of *Mtb*-derived constituents as diagnostic biomarkers have been cross-sectional (46), so it is not known how long specific *Mtb*-derived metabolites persist in the infected host. If it is confirmed that metabolites continue to circulate for weeks despite effective antibiotic therapy then this may limit their use as a test of treatment response or cure. Longer-term studies on the kinetics of detection of the metabolites we targeted would be of interest in this regard.

Of the four metabolites identified as potential diagnostic biomarkers in the discovery analysis, only Lyso-PIM1 was significantly different between patients with drug susceptible TB and patients with MDR TB. We found that Lyso-PIM1 was significantly increased in patients with active TB due to a drug susceptible strain of *Mtb*. This finding is contrary to our *pre hoc* hypothesis that circulating *Mtb* cell envelope metabolites may be indicative of ongoing organism replication and may therefore be higher in patients with MDR TB, who were not receiving effective antibiotic therapy.

Similarly, of the 30 *Mtb* lipid metabolites significantly different between drug susceptible and MDR TB patients, 28 were increased in patients with drug susceptible TB; an effect that persisted throughout the first 8 weeks of treatment. This suggests that in patients undergoing effective TB treatment, increased concentrations of *Mtb* lipid metabolites may reflect increased organism killing rather than continued replication, although this conclusion remains speculative. All patients received multiple cell wall-active antibiotics for treatment and in patients with drug susceptible disease this would have resulted in increased organism killing relative to those with MDR TB. Patients within one week of starting treatment were eligible to be included in the study, and nearly all patients received at least three days of antibiotics prior to study enrollment. With our study design, it is not possible to know whether these differences were present prior to antibiotic initiation. Given these *Mtb* lipids were already increased in drug susceptible TB patients in the baseline plasma sample, it will be important for future studies to examine *Mtb* metabolite concentrations prior to starting antibiotics (a better choice for time zero). This will help determine whether higher concentrations of *Mtb* lipids in drug sensitive TB patients are an early consequence of effective antibiotic therapy.

Another possible explanation for our data is that the cell envelope of MDR strains of *Mtb* is structurally altered and therefore produces lower concentrations of certain lipid metabolites compared to drug susceptible strains. In cell culture studies, Lahiri et al. found *Mtb* strains resistant to rifampin had altered concentrations of several *Mtb* lipids in the cell envelope (50). One of the lipid species present in significantly lower concentrations in the cell envelope of rifampin-resistant *Mtb* strains was acylated sulfoglycolipid, which was also found to be decreased in the plasma of MDR TB patients in the current study. These authors hypothesized that alterations in gene transcription caused by mutations conferring rifampin resistance may result in *Mtb* cell wall remodeling. If such cell wall lipid remodeling can be detected through plasma metabolomic profiling in patients with active TB, plasma HRM may be a strategy to explore identification of MDR TB at the time of initial TB disease diagnosis.

This study has several strengths. We used a unique analytic approach coupling robust plasma HRM with rigorous quality control (51) and an open-access published *Mtb* cell wall metabolite database (32) to identify *Mtb* cell envelope metabolites in human plasma. The Georgian household contacts were genetically similar to the active TB cases and therefore differences in *Mtb* metabolites between groups are unlikely to reflect genetically mediated host differences in metabolism of such metabolites. All active TB cases were enrolled in a randomized clinical trial, with state-of-the at diagnostics and TB treatment in a WHO-accredited national TB treatment center; plasma samples and clinical data were collected in a systematic fashion, with close follow up throughout the study period.

This study is subject to several limitations. The sample size for this study was modest with only 17 pulmonary TB cases and 16 household contacts included in the cross-sectional biomarker discovery analysis. Due to the small sample size, the four *Mtb* biomarkers found to distinguish patients with active TB from their contacts have not been validated on an independent cohort of active TB patients and controls without active TB disease; this awaits future study. In addition, the small sample size limited the statistical power to control for multiple covariates. However, only smoking status was significantly different between TB patients and household contacts and this is unlikely to account for differences in circulating *Mtb* metabolites.

Our study was conducted at a single center in the country of Georgia where there is a relatively homogenous (Caucasian) population. Further work is needed to investigate the presence of these metabolites in other patient populations. Generalizability may also be limited for active TB patients with a negative AFB smear and/or culture, given nearly all patients enrolled in this study had a positive result for both. Additionally, the detection range of the high-resolution mass spectrometer was not optimized for all *Mtb* cell wall lipids.

Certain *Mtb* cell wall lipids have a larger mass and therefore require a detection range of up to 3,000 Daltons for identification. Finally, because no purified standards are available for the identified metabolites at present, chemical confirmation of *Mtb* lipid suspects using tandem MS/MS of authentic standards could not be performed.

In future studies, it will be important to validate the *Mtb* lipid metabolites identified in this study in a larger population of active TB patients and using purified authentic standards. If these metabolites can differentiate active TB cases from controls without active TB disease or patients with LTBI in an independent population then this will increase the likelihood they can be developed as a diagnostic biosignature for active TB disease. It will also be essential to further refine the plasma HRM assay to more comprehensively detect less polar *Mtb*-derived lipid metabolites, with LC-HRMS in negative electrospray ionization mode and over a greater m/z coverage range. Further studies are also needed to determine when in the course of infection *Mtb*-derived lipids become detectable in plasma of infected patients, how long they continue to circulate after appropriate treatment and how they may relate to long-term clinical outcomes.

In conclusion, we found that among several *Mtb* lipid metabolite matches detected, Lyso-PIM1, Lyso-PI, PG and Ac1PIM1 were the most significantly elevated in the plasma of patients with active TB disease compared to their asymptomatic household contacts. These metabolites provided excellent classification accuracy for active TB (AUC 0.95). We further showed that certain specific *Mtb* lipid metabolites were increased over the course of treatment in patients with drug susceptible TB compared to those infected with MDR strains. These findings suggest plasma HRM may hold significant promise for discovery of potential biomarkers to detect patients with active TB disease. If confirmed in large studies that include diverse patient populations, further development of the most promising biomarkers could be used to develop a rapid point of care test for diagnosis and treatment response.

References

1. WHO. Global tuberculosis report 2016: World Health Organization; 2016.
2. Lienhardt C, et al. Translational Research for Tuberculosis Elimination: Priorities, Challenges, and Actions. 2016 Mar 2;13(3):e1001965.
3. Johnson CH, et al. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016 Jul;17(7):451-9.
4. Walker DI, et al. Population Screening for Biological and Environmental Properties of the Human Metabolic Phenotype: Implications for Personalized Medicine. In: Metabolic Phenotyping in Personalized and Public Healthcare: Elsevier Inc.; 2016. p. 167-211.
5. Tukvadze N, et al. High-dose vitamin D3 in adults with pulmonary tuberculosis: a double-blind randomized controlled trial. Am J Clin Nutr. 2015 Nov;102(5):1059-69.
6. Pai M, et al. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. J Infect Dis. 2015 Apr 1;211 Suppl 2:S21-8.
7. Perkins MD, et al. Diagnostic testing in the control of tuberculosis. Bull World Health Organ. 2002; 80:512-3.
8. Sharma SK, et al. Relevance of latent TB infection in areas of high TB prevalence. Chest. 2012 Sep;142(3):761-73.
9. Jensen PA, et al. Guidelines for preventing the transmission of Mycobacterium tuberculosis in health-care settings, 2005. MMWR Recomm Rep. 2005 Dec 30;54(RR-17):1-141.
10. Sepkowitz KA. How Contagious Is Tuberculosis? Clin Infect Dis. 1996 Nov;23(5):954-62.
11. Davis JL, et al. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis. 2013 Feb;13(2):147-54.
12. Steingart KR, et al. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis. 2006 Oct;6(10):664-74.
13. Getahun H, et al. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. Lancet. 2007 Jun 16;369(9578):2042-9.
14. Puri L, et al. Xpert MTB/RIF for tuberculosis testing: access and price in highly privatised health markets. Lancet Glob Health. 2016 Feb;4(2):e94-5.
15. McNerney R, et al. Towards a point-of-care test for active tuberculosis: obstacles and opportunities. Nat Rev Microbiol. 2011 Mar;9(3):204-13.
16. Yerlikaya S, et al. A tuberculosis biomarker database: the key to novel TB diagnostics. Int J Infect Dis. 2017 Mar;56:253-257.
17. Weiner J, et al. High-throughput and computational approaches for diagnostic and prognostic host tuberculosis biomarkers. Int J Infect Dis. 2017 Mar;56:258-262.
18. Steingart KR, et al. Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: an updated systematic review and meta-analysis. PLoS Med. 2011 Aug;8(8):e1001062.
19. Pai M, et al. Systematic Review: T-Cell–based Assays for the Diagnosis of Latent Tuberculosis Infection: An Update. Ann Intern Med. 2008 Aug 5;149(3):177-84.
20. Shah M, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. Cochrane Database Syst Rev. 2016 May 10;(5):CD011420.

21. WHO. Treatment of Tuberculosis Guidelines. World Health Organization; 2010.
22. Gillespie SH, et al. Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. N Engl J Med. 2014 Oct 23;371(17):1577-87.
23. Frediani JK, et al. Plasma metabolomics in human pulmonary tuberculosis disease: a pilot study. PLoS One. 2014 Oct 15;9(10):e108854.
24. Collins JM, et al. Metabolomics and Mycobacterial Disease: Don't Forget the Bioinformatics. Ann Am Thorac Soc. 2016 Jan;13(1):141-2.
25. Go YM, et al. Reference Standardization for Mass Spectrometry and High-Resolution Metabolomics Applications to Exposome Research. Toxicol Sci. 2015 Dec;148(2):531-43.
26. Lau SK, et al. Metabolomic Profiling of Plasma from Patients with Tuberculosis by Use of Untargeted Mass Spectrometry Reveals Novel Biomarkers for Diagnosis. J Clin Microbiol. 2015 Dec;53(12):3750-9.
27. Weiner J 3rd, et al. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. PLoS One. 2012;7(7):e40221.
28. da Silva RR, et al. Illuminating the dark matter in metabolomics. Proc Natl Acad Sci U S A. 2015 Oct 13;112(41):12549-50.
29. Pan SJ, et al. Biomarkers for Tuberculosis Based on Secreted, Species-Specific, Bacterial Small Molecules. J Infect Dis. 2015 Dec 1;212(11):1827-34.
30. Shui G, et al. Mycolic acids as diagnostic markers for tuberculosis case detection in humans and drug efficacy in mice. EMBO Mol Med 2012; 4:27-37.
31. Mahapatra S, et al. A metabolic biosignature of early response to anti-tuberculosis treatment. BMC Infect Dis. 2014 Jan 31;14:53.
32. Sartain MJ, et al. Lipidomic analyses of Mycobacterium tuberculosis based on accurate mass measurements and the novel "Mtb LipidDB". J Lipid Res. 2011 May;52(5):861-72.
33. Layre E, et al. A comparative lipidomics platform for chemotaxonomic analysis of Mycobacterium tuberculosis. Chem Biol. 2011 Dec 23;18(12):1537-49.
34. Lomtadze N, et al. Prevalence and risk factors for multidrug-resistant tuberculosis in the Republic of Georgia: a population-based study. Int J Tuberc Lung Dis. 2009 Jan;13(1):68-73.
35. Tukvadze N, et al. Use of a molecular diagnostic test in AFB smear positive tuberculosis suspects greatly reduces time to detection of multidrug resistant tuberculosis. PLoS One. 2012;7(2):e31563.
36. Soltow QA, et al. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. Metabolomics. 2013 Mar;9(1 Suppl):S132-S143.
37. Yu T, et al. apLCMS--adaptive processing of high-resolution LC/MS data. Bioinformatics. 2009 Aug 1;25(15):1930-6.
38. Uppal K, et al. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. BMC Bioinformatics. 2013 Jan 16;14:15.
39. Bade R, et al. Suspect screening of large numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time prediction and high resolution mass spectrometry data analysis. Sci Total Environ. 2015 Dec 15;538:934-41.

40. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.
41. Hochberg Y, et al. More powerful procedures for multiple significance testing. Stat Med. 1990 Jul;9(7):811-8.
42. Bates D, et al. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw. 2015; 67: 48.
43. Ogata H, et al. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 1999 Jan 1;27(1):29-34.
44. Smith CA, et al. METLIN: a metabolite mass spectral database. Ther Drug Monit. 2005 Dec;27(6):747-51.
45. Wishart DS, et al. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007 Jan;35(Database issue):D521-6.
46. Haas CT, et al. Diagnostic 'omics' for active tuberculosis. BMC Med. 2016 Mar 23;14:37.
47. Mishra AK, et al. Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in Mycobacterium tuberculosis physiology and host-pathogen interaction. FEMS Microbiol Rev. 2011 Nov;35(6):1126-57.
48. Esmail H, et al. Characterization of progressive HIV-associated tuberculosis using 2-deoxy-2-[18F]fluoro-D-glucose positron emission and computed tomography. Nat Med. 2016 Oct;22(10):1090-1093.
49. Crellin PK, et al. (2013). Metabolism of Plasma Membrane Lipids in Mycobacteria and Corynebacteria, Lipid Metabolism, Prof. Rodrigo Valenzuela Baez (Ed.), InTech, Available from: https://www.intechopen.com/books/lipid-metabolism/metabolism-of-plasma-membrane-lipids-in-mycobacteria-and-corynebacteria
50. Lahiri N, et al. Rifampin Resistance Mutations are Associated with Broad Chemical Remodeling of Mycobacterium tuberculosis. J Biol Chem. 2016 Jul 1;291(27):14248-56.
51. Jones DP. Sequencing the exposome: A call to action. Toxicol Rep. 2016;3:29-45.

Table 1 – Characteristics of patients diagnosed with pulmonary tuberculosis and their asymptomatic household contacts

	Active TB cases (n=17)	Household Contacts (n=16)	P-value†
Age [years; mean (SD)]*	33.4 (11.8)	38.6 (10.2)	0.18
Male Sex, n (%)	10 (59%)	6 (38%)	0.22
Tobacco Use	13 (41%)	6 (38%)	0.02
AFB Sputum Smear Positive‡	16 (94%)	N/A	N/A
Cavitation on Chest X-ray	2 (12%)	N/A	N/A
Multidrug-resistant (MDR) TB§	3 (18%)	N/A	N/A

*SD = standard deviation
† A two-tailed t test was used for comparisons of continuous data and a two-tailed Fisher exact test was used for categorical data
‡ Sputum smear obtained at study enrollment
§ Multidrug resistance was defined as resistant to both isoniazid and rifampin by drug susceptibility testing

Table 2 –*Mycobacterium tuberculosis* lipid metabolites significantly increased in patients with active TB infection compared to household contacts

Metabolite	m/z*	Retention Time (seconds)	Mean Intensity in Persons with active TB (n=17)†	Mean Intensity in household contacts (n=16)	P-value‡
Ac1PIM1					
R1CO2H+R2CO2H +R3CO2H=56:1	1321.9177	427.25	13.08	2.81	<0.001
Ac1PIM2					
R1CO2H+R2CO2H +R3CO2H=46:0	1367.8083	524.15	10.54	5.45	0.05
R1CO2H+R2CO2H +R3CO2H=49:3	1403.8277	489.57	16.60	11.98	0.01
CL					
2R1CO2H+ 2R2CO2H=68:3	1385.9836	308.30	10.53	2.99	0.005
Lyso-PE					
R1CO2H=18:0	464.3102	449.33	20.38	19.37	0.02
Lyso-PG					
RCO2H=17:1	497.2916	318.92	8.26	3.39	0.04
Lyso-PI					
RCO2H=18:0	583.3219	363.72	10.75	0.79	<0.001
Lyso-PIM1					
RCO2H=17:0	749.3714	279.43	10.75	1.08	<0.001
MG					
RCO2H=18:0	341.3026	510.52	18.58	17.97	0.05
PE					
R1CO2H+R2CO2H =37:2	758.5627	562.14	19.15	8.67	0.006
PG					
R1CO2H+R2CO2H =34:1	749.5284	558.02	12.56	3.55	<0.001
PIM1					
R1CO2H+R2CO2H =36:3	1005.5899	500.06	17.01	16.32	0.02
PIM2					
R1CO2H+R2CO2H =34:4	1155.6035	290.55	7.85	2.80	0.04
R1CO2H+R2CO2H =36:0	1173.6807	352.47	11.67	4.35	0.002

* mass to charge ratio of metabolite

† semi-quantitative measure of metabolite concentration; all intensity measurements were log2 transformed

‡ p-value comparing mean metabolite intensity between groups by LIMMA

Table 3 – Characteristics of patients with drug susceptible and multidrug resistant pulmonary tuberculosis followed for response to treatment

	Drug Sensitive TB (n=44)	MDR TB (n=17)†	P-value‡
Age – [years; mean (SD)]*	31.8 (11.6)	31.6 (9.7)	0.96
Male sex – n (%)	26 (59.1)	8 (47.1)	0.40
Current Smoker – n (%)	29 (65.9)	8 (47.1)	0.18
HIV Positive – n (%)	1 (2.3)	0 (0)	1.00
Diabetes – n (%)	3 (6.8)	1 (5.9)	1.00
Cavitary disease – n (%)	7 (15.9)	4 (23.5)	0.48
Extrapulmonary disease – n (%)	2 (4.5)	0 (0)	1.00
Sputum culture negative at 4 weeks – n (%)	23 (54.8)	6 (35.3)	0.13
Sputum culture negative at 8 weeks – n (%)	39 (92.9)	8 (47.1)	<0.001

*SD = standard deviation

† Multidrug resistance was defined as resistant to both isoniazid and rifampin

‡ A two-tailed t test was used for comparisons of continuous data and a two-tailed Fisher exact test was used for categorical data

Table 4 –*Mycobacterium tuberculosis* lipid metabolites significantly increased in patients with drug susceptible TB relative to those with multidrug resistant TB

Metabolite	m/z*	Retention Time (seconds)	Mean increase in intensity in drug sensitive TB cases†	P-value‡
PIM5				
R1CO2H+R2CO2H=34:1	1669.8041	366.57	5.50	<0.001
Ac2PIM2				
R1CO2H+R2CO2H+ R3CO2H+R4CO2H=64:4	1626.0137	486.43	3.49	0.001
R1CO2H+R2CO2H+ R3CO2H+R4CO2H=68:5	1658.0766	446.02	1.86	0.02
Ac1PIM3				
R1CO2H+R2CO2H+ R3CO2H=53:0,R4=H	1605.9845	432.54	3.38	0.002
R1CO2H+R2CO2H+ R3CO2H=55:3,R4=H	1609.9641	429.91	3.17	0.03
Trehalose monomycolate				
Alpha-MA, C1616.4591434	1598.4601	432.65	3.45	0.003
Alpha mycolic acid				
C75	1078.1288	431.63	1.07	0.006
Lyso-PIM1				
RCO2H=19:2	773.3737	371.48	3.81	0.008
RCO2H=17:0	731.3646	371.98	3.98	0.02
PI				
R1CO2H+R2CO2H=36:2	863.5653	474.82	3.81	0.01
R1CO2H+R2CO2H=35:0	853.5834	521.17	3.15	0.02
PIM5				
R1CO2H+R2CO2H=34:1	1647.8177	372.66	3.71	0.01
R1CO2H+R2CO2H=34:2	1667.7949	371.06	4.33	0.02
Ac1PIM4				
R1CO2H+R2CO2H+ R3CO2H=45:0,R4=H	1637.9239	432.82	2.70	0.02
Keto mycolic acid				
C82	1210.2397	432.92	0.75	0.02
Ac1PIM2				
R1CO2H+R2CO2H+ R3CO2H=57:1,R4=H	1479.9858	486.72	2.96	0.02
PIM1				
R1CO2H+R2CO2H=36:3	1005.5899	500.06	17.01	0.02
PIM2				
R1CO2H+R2CO2H=34:4	1155.6035	290.55	7.85	0.04
R1CO2H+R2CO2H=36:0	1173.6807	352.47	11.67	0.002

* mass to charge ratio of metabolite
† semi-quantitative measure of metabolite concentration
‡ p-value of MDR status in mixed effects model using a likelihood ratio test

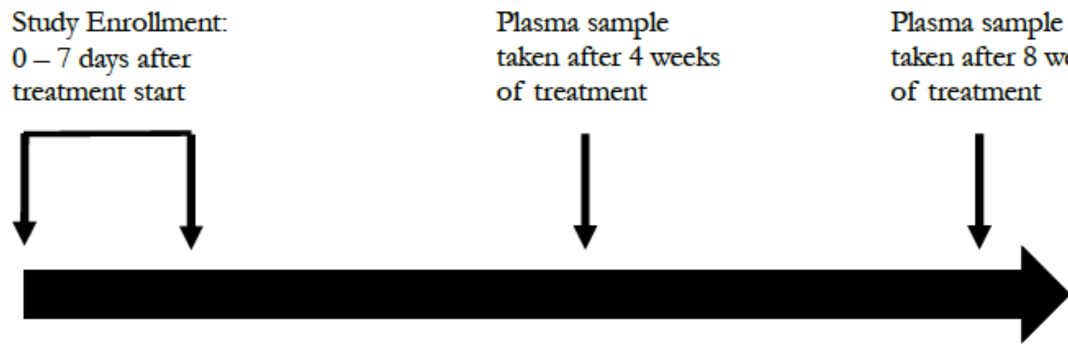

Figure 1 - Study timeline for measurement of plasma metabolomics during treatment. Per the protocol of the original study, sputum and peripheral blood plasma samples from active TB cases were collected at baseline after 2, 4, 6, 8, 12 and 16 weeks of treatment (5).

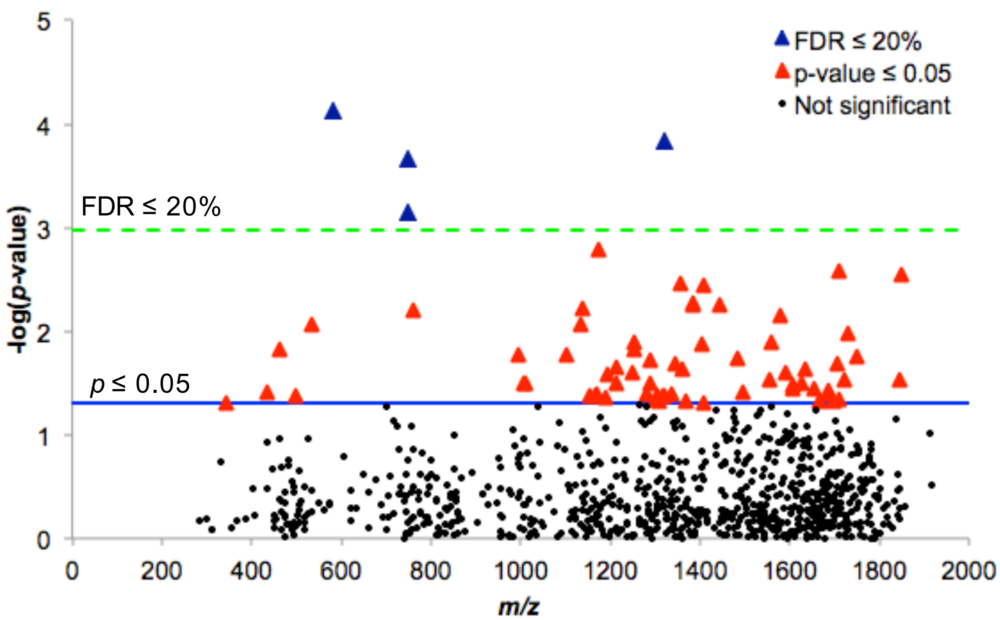

Figure 2 - Differentially Expressed Metabolites. Type 2 Manhattan plot of cross-sectional comparison of targeted *Mycobacterium tuberculosis (Mtb)* cell wall lipid metabolites in plasma of 16 household contacts without active TB disease and 17 patients with active pulmonary TB disease. A total of 867 features were identified as *Mtb*-associated lipids based on accurate m/z matches. Sixty-nine *Mtb* lipids were significant at raw $p \leq 0.05$ (red triangles) and four were significant using a false discovery rate (FDR) correction of 20% (blue triangles). Black dots below the blue horizontal line represent detected *Mtb* lipid suspects not significantly different between groups. The negative $\log_{10}$ statistical p-value of metabolites between the two groups are shown on the y-axis as a function of metabolite m/z, (x-axis).

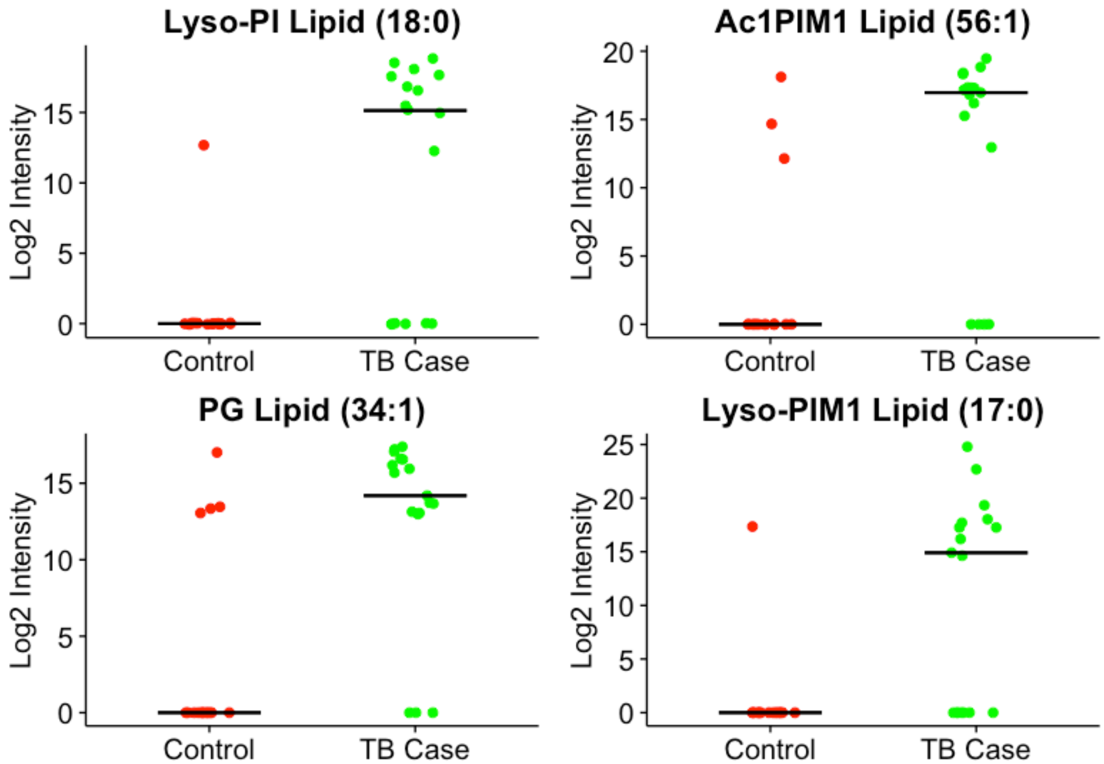

Figure 3 - Four most significant *Mycobacterium tuberculosis* lipid matches. Putative identification and log2 intensity values of the *Mtb* lipid metabolites phosphatidylglycerol (PG), monoacylglycerophosphoinositol (Lyso-PI), monoacylated diacylglycerophosphoinositolmonomannoside (Ac1PIM1), and Monoacylglycerophosphoinositolmonomannoside (Lyso-PIM1) (each raw $p<0.001$; active TB cases vs household contacts).

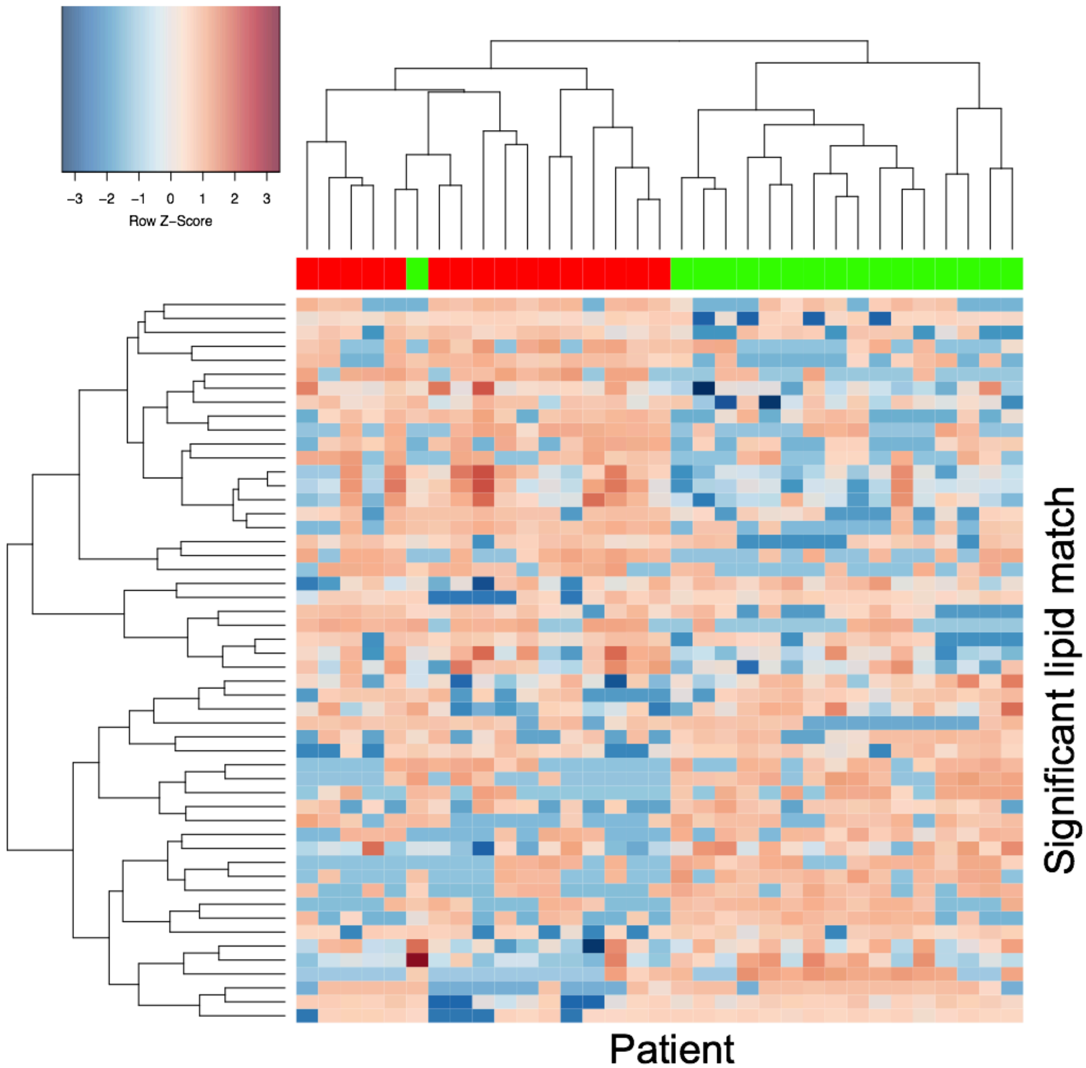

Figure 4 - Hierarchical Cluster Analysis. Using the 69 differentially expressed *Mtb* lipid suspects, the two-way HCA shows clear separation between the 17 active TB cases (green) and the 16 household contacts (red) shown along the upper x-axis. The 69 Mtb lipid metabolite matches are depicted in rows along the y-axis, with the fold-change increase or decrease between cases and controls for each metabolite within individual subjects color-coded (indicated by the color key).

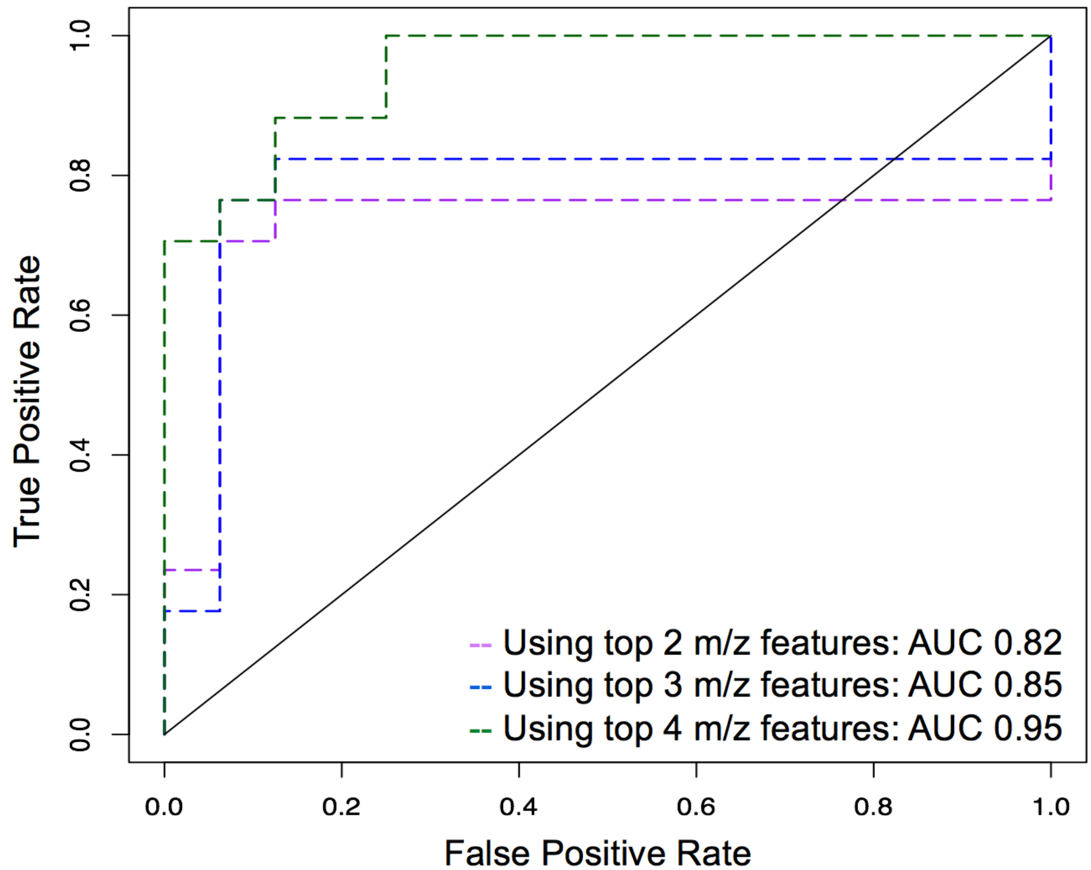

Figure 5 - Receiver Operating Characteristic Curve. The four *Mtb* metabolites PG, Lyso-PI, Ac1PIM1 and Lyso-PIM1, elevated in patients with active pulmonary TB relative to household controls, provided excellent classification accuracy for active TB (AUC of 0.95).

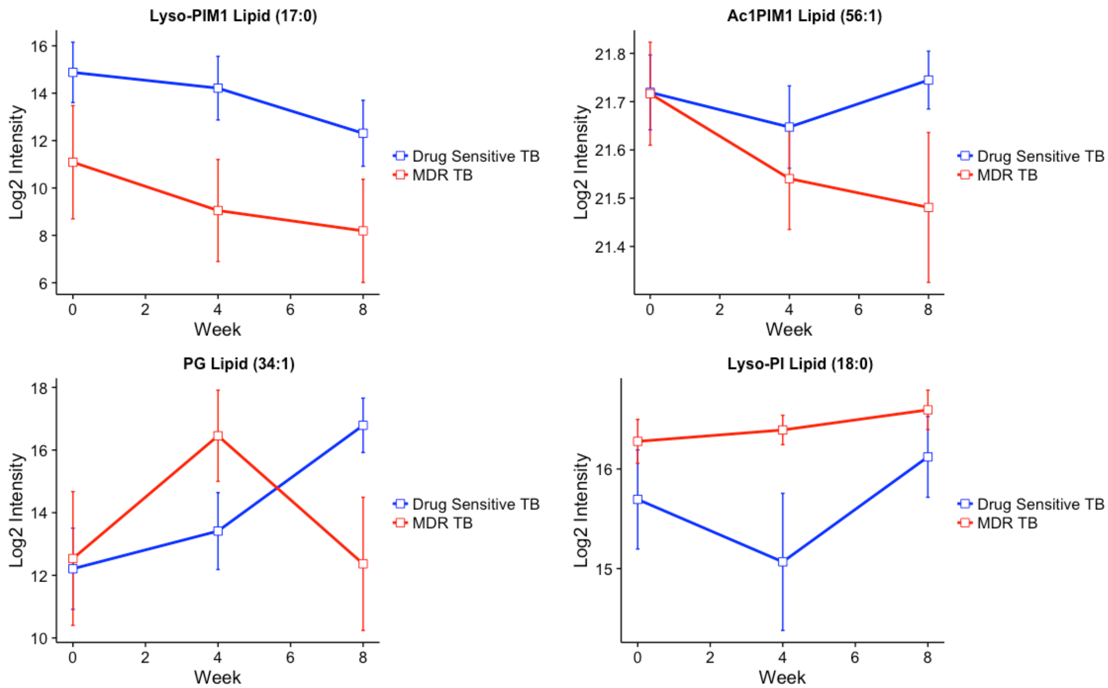

Figure 6 - *Mycobacterium tuberculosis* lipid metabolites during treatment. For the *Mtb* lipids significantly associated with active TB disease, all remained detectable during the first 8 weeks of treatment, regardless of MDR status. Lyso-PIM1 was significantly elevated in patients with drug sensitive TB at all time points.