

RESEARCH ARTICLE

High-resolution plasma metabolomics analysis to detect *Mycobacterium tuberculosis*-associated metabolites that distinguish active pulmonary tuberculosis in humans

Jeffrey M. Collins¹*, Douglas I. Walker², Dean P. Jones³, Nestani Tukvadze⁴, Ken H. Liu³, ViLinh T. Tran³, Karan Uppal³, Jennifer K. Frediani⁵, Kirk A. Easley⁶, Neeta Shenvi⁶, Manoj Khadka⁷, Eric A. Ortlund⁷, Russell R. Kempker¹, Henry M. Blumberg¹, Thomas R. Ziegler^{8,9}



1 Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia, United States of America, **2** Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America, **3** Clinical Biomarkers Laboratory, Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia, United States of America, **4** National Center for Tuberculosis and Lung Disease, Tbilisi, Georgia, **5** Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, Georgia, United States of America, **6** Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America, **7** Emory Integrated Lipidomics Core, Department of Biochemistry, Emory University School of Medicine, Atlanta, Georgia, United States of America, **8** Division of Endocrinology, Metabolism and Lipids and Center for Clinical and Molecular Nutrition, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia, United States of America, **9** Section of Endocrinology, Atlanta Veterans Affairs Medical Center, Atlanta, Georgia, United States of America

* These authors contributed equally to this work.

* jmcoll4@emory.edu

OPEN ACCESS

Citation: Collins JM, Walker DI, Jones DP, Tukvadze N, Liu KH, Tran VT, et al. (2018) High-resolution plasma metabolomics analysis to detect *Mycobacterium tuberculosis*-associated metabolites that distinguish active pulmonary tuberculosis in humans. PLoS ONE 13(10): e0205398. <https://doi.org/10.1371/journal.pone.0205398>

Editor: Selvakumar Subbian, Rutgers Biomedical and Health Sciences, UNITED STATES

Received: January 25, 2018

Accepted: September 25, 2018

Published: October 11, 2018

Copyright: © 2018 Collins et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported, in part, by grants from the National Institutes of Health: [R21 AI130918 to TRZ, DPJ, DIW, RRR, HMB], [T32 AI074492 to JMC], [K24 DK096574 to TRZ], [D43 TW007124 to HMB], [K23 AI103044 to RRR], [S10 OD018006 to DPJ, DIW], [P30 ES019776 to DPJ,

Abstract

Introduction

Pulmonary tuberculosis (TB) is a major worldwide health problem that lacks robust blood-based biomarkers for detection of active disease. High-resolution metabolomics (HRM) is an innovative method to discover low-abundance metabolites as putative blood biomarkers to detect TB disease, including those known to be produced by the causative organism, *Mycobacterium tuberculosis* (*Mtb*).

Methods

We used HRM profiling to measure the plasma metabolome for 17 adults with active pulmonary TB disease and 16 of their household contacts without active TB. We used a suspect screening approach to identify metabolites previously described in cell culture studies of *Mtb* based on retention time and accurate mass matches.

Results

The association of relative metabolite abundance in active TB disease subjects compared to their household contacts predicted three *Mtb*-associated metabolites that were

TRZ, DIW], [UL1 TR000454 (Atlanta Clinical and Translational Science Institute)], the Emory Medical Care Foundation [2016/06-0002 to RRK] and the Emory University Global Health Institute. The funding sources had no involvement in the study design; collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to submit the article for publication.

Competing interests: The authors have declared that no competing interests exist.

significantly increased in the active TB patients based on accurate mass matches: phosphatidylglycerol (PG) (16:0_18:1), lysophosphatidylinositol (Lyso-PI) (18:0) and acylphosphatidylinositol mannoside (Ac1PIM1) (56:1) ($p < 0.001$ for all). These three metabolites provided excellent classification accuracy for active TB disease (AUC = 0.97). Ion dissociation spectra (tandem MS/MS) supported the identification of PG (16:0_18:1) and Lyso-PI (18:0) in the plasma of patients with active TB disease, though the identity of Ac1PIM1 could not be definitively confirmed.

Conclusions

Presence of the *Mtb*-associated lipid metabolites PG (16:0_18:1) and Lyso-PI (18:0) in plasma accurately identified patients with active TB disease. Consistency of *in vitro* and *in vivo* data suggests suitability for exploring these in future studies for possible development as TB disease biomarkers.

Introduction

Tuberculosis (TB) is a major global health problem and now the leading cause of death due to an infectious disease [1]. Rapid and accurate detection of active TB disease remains a major challenge in global control efforts, with less than two thirds of estimated TB cases diagnosed in 2016 [1]. The poor performance characteristics of available diagnostic tests for pulmonary TB, especially smear microscopy, which is frequently used in low- and middle-income countries, is a major contributor to suboptimal case detection and diagnostic delays [2]. Acid-fast bacilli (AFB) culture is the gold standard for diagnosis, but can take 2–6 weeks for definitive results and the infrastructure required for culture is often unavailable in resource-limited settings [3]. Furthermore, all tests require analysis of sputum specimens, which many adult patients and most children cannot adequately produce. There is an urgent need for point-of-care diagnostic tests for active TB which can be performed on peripheral blood or urine samples [3]. The development of such tests will initially require identification of new putative biomarkers in the peripheral blood associated with active TB disease [4, 5].

Recently developed high-resolution metabolomics (HRM) methods utilize liquid chromatography and ultra-high-resolution mass spectrometry (LC-MS), coupled with advanced methods in data extraction and bioinformatics, to detect tens of thousands of metabolites in plasma and other biosamples [6–8]. The high sensitivity of HRM allows for simultaneous detection and semi-quantitative measurement of multiple low-abundance metabolites, including those potentially derived from infecting microorganisms such as *Mycobacterium tuberculosis* (*Mtb*) [7, 8]. Yet the high number of chemical features detected creates challenges for comprehensive metabolite annotation. It is now apparent that searchable metabolite databases contain only a small fraction of various metabolites derived from humans, microorganisms and environmental chemicals; thus, many features detected by HRM are not identified and remain unknown chemicals [7, 9, 10]. Recent studies, including our own, have focused on untargeted (discovery) evaluation of changes in systemic endogenously-derived metabolites associated with active TB disease compared to controls without active TB [7, 11, 12]. However, the current sensitivity of HRM methods now enables the characterization of low-abundance metabolites in the human host, some of which may be pathogen-derived with the potential to provide new insights into host-pathogen interactions and pathophysiology.

In our preliminary studies using plasma HRM, we used available open-access human metabolite databases for metabolite annotation [13–15], with nearly half of the 61 metabolites differentiating TB cases from household contacts lacking matches to known metabolites based on accurate mass/charge (m/z) ratio and predicted adduct forms [7]. Because existing databases contain few metabolites derived from *Mtb*, we hypothesized that some of these “unknown” metabolites may be derived from *Mtb* in the infected host. Recently, using a combined LC-MS and computational approach, databases of metabolites derived from *Mtb* have been developed [16, 17]. We used these databases in a suspect screening approach to re-analyze the plasma HRM data from our previous investigation [7] for evidence of *Mtb*-associated metabolites in plasma that differentiate patients with newly diagnosed active pulmonary TB from asymptomatic household contacts without active TB.

Methods

Study design

We performed a cross-sectional analysis of plasma HRM profiles for 17 patients with active TB disease selected from a randomized, controlled trial of high-dose cholecalciferol treatment of patients with pulmonary TB conducted in the country of Georgia (clinicaltrials.gov identifier NCT00918086) [18]. Inclusion criteria for patients included age ≥ 18 years, newly diagnosed, symptomatic active TB disease, suggested by a positive AFB sputum smear and confirmed by semi-quantitative sputum culture for *Mtb* (performed at the Georgian National TB Reference Laboratory (NRL) [18]). Patients with other localizing symptoms were evaluated for the presence of extrapulmonary TB. Plasma for HRM was obtained from eligible subjects within 7 days of initiating anti-TB drug therapy with first-line anti-TB drugs (isoniazid, rifampicin, pyrazinamide and ethambutol) [18]. Drug susceptibility testing (DST) was performed on all persons with pulmonary TB using the absolute concentration method. Patients infected with *Mtb* isolates resistant to isoniazid and rifampin were considered to have multidrug resistant (MDR)-TB.

The control group consisted of 16 household contacts enrolled from persons accompanying patients to the TB treatment facility at the time of study enrollment (one household contact was excluded from analysis for technical reasons). To be eligible for enrollment, household contacts were required to have both a negative screen for clinical symptoms suggestive of active TB or other acute illness and a negative AFB sputum smear and culture. Household contacts were not evaluated with a chest radiograph and were not tested for latent TB infection (LTBI).

Plasma sample collection

Peripheral blood samples were obtained by venipuncture from all subjects with TB disease and the household contacts upon entry into the study [18]. Blood was collected in ethylenediaminetetraacetic acid (EDTA)-containing tubes and centrifuged; isolated plasma was immediately frozen and stored at -80°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 -80°C prior to metabolomics analysis.

Plasma metabolomics analysis

Thawed plasma (65 μL) was treated with 130 μL acetonitrile (2:1, v/v) containing an internal isotopic standard mixture (3.5 μL /sample), as previously described [7,19]. The internal standard mix for quality control consisted of 14 stable isotopic chemicals covering a broad range

of small molecules [19]. Samples were mixed and placed on ice for 30 min prior to centrifugation to remove protein. The resulting supernatant was transferred to low-volume autosampler vials maintained at 4°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) with formic acid/acetonitrile gradient [7]. The high-resolution mass spectrometer was operated in positive electrospray ionization mode over scan range of 85 to 2000 *m/z* and stored as .Raw files [6, 7]. Data were extracted and aligned using apLCMS [20] and xMSanalyzer [21] with each feature defined by specific *m/z* value, retention time and integrated ion intensity. Non-zero ion intensity values from technical replicates were averaged to create a single mean intensity value for each participant.

Metabolite identification

A suspect screening approach was used to identify potential *Mtb* lipid metabolites present in the plasma samples from subjects with active TB disease and household contacts without TB. A previously published *Mtb* lipid library, “*Mtb* LipidDB” was used to annotate *Mtb* lipid candidates based on accurate mass matches to specific *m/z* in the *Mtb* library [16]. The database search included three metabolite (M) adducts for each lipid species (*m/z*): plus hydrogen (H), plus sodium (Na), and minus water (H₂O)/plus hydrogen, respectively (M+H, M+Na, and M-H₂O+H), with a mass accuracy threshold of 10 parts-per-million (ppm).

For those *Mtb* lipid metabolite suspects found to be significantly increased in patients with active TB disease relative to household contacts, we attempted to verify their chemical identities using ion dissociation methods (MS/MS). To maximize the ability to detect interpretable fragmentation spectra for lipid species of interest, additional plasma samples were extracted by adding 400 µl methanol followed by addition of a methyl *tert*-butyl ether (MTBE):methanol (1:1 v/v) mixture, and finally with 500 µl of 100% MTBE. All extract fractions were pooled and dried under a gentle stream of nitrogen. The dried extracts were resoluted in 500 µl of a chloroform:methanol (1:1 v/v) mixture followed by direct infusion analysis using a AB SCIEX QTRAP 5500 triple quadrupole mass spectrometer in product ion scan mode. MS/MS fragmentation of [M-H] adducts was performed in negative ionization mode. The MS/MS spectra were then investigated to find the lipid-specific fragments to confirm the chemical identity of the metabolites of interest.

Data analysis

All analyses were performed in R version 3.3.0. Comparisons of descriptive statistics and clinical characteristics between subjects with pulmonary TB and their household contacts were performed using a Wilcoxon rank sum test for continuous variables and a Fisher exact test for categorical variables. The intensity profiles for *Mtb* lipid suspects were log₂ transformed and tested for association with disease status using linear model for microarray data (LIMMA) analysis [22]. A Benjamini-Hochberg false discovery rate (FDR) of 20% was then applied to adjust for multiple comparisons and to identify those metabolites of greatest interest [23]. Logistic regression analysis was used to construct receiver operating characteristic (ROC) curves to determine the classification accuracy of the metabolites with the greatest -log P value.

IRB approval

This study was approved by the Institutional Review Board (IRB) of Emory University (Atlanta, GA, USA) and the Georgian National Center for Tuberculosis and Lung Disease (NCTLD) Ethics Committee (Tbilisi, Georgia). All subjects provided written informed consent.

Results

Participants

Sex did not differ significantly in the 17 active TB cases compared to the 16 household contacts (Table 1). Those with active TB were more likely to smoke compared to the household contacts (76% vs 38%, $p = 0.02$) and had a lower median age (27 vs 35 years, $p = 0.05$). None of the patients with pulmonary TB had evidence of extrapulmonary disease. Three (18%) of the TB patients had MDR-TB and two (12%) had evidence of cavitation on chest x-ray. None of the subjects had HIV co-infection.

High-resolution metabolomics profiling results

Following data extraction and alignment with aPLCMS and xMSanalyzer, 33,262 mass spectral features, defined by accurate mass m/z , chromatographic retention time and intensity were detected. To maximize detection of low-abundance metabolites, three technical replicates were performed, and we removed features with a coefficient of variation (CV) $\geq 100\%$ between technical replicates, resulting in 32,975 remaining features. The m/z values were then tested for accurate mass matches to the 1,696 unique lipid species present in *Mtb* LipidDB with adduct masses within the instrument scan range (85–2000 da) [16]. Using the criteria above, 2,470 matches to *Mtb* lipid suspects were detected by HRM (including M+H, M+Na, and M-H₂O +H adducts of each specific m/z species). To target reliably detectable, non-polar ions consistent with *Mtb* lipids, we conservatively removed features exhibiting retention time ≤ 3 minutes and those with $\geq 50\%$ missing values in both individuals with active TB disease and the household contacts, resulting in 867 remaining *Mtb* metabolite suspects.

Using LIMMA to test the association of \log_2 transformed intensity values of *Mtb* metabolite matches with active TB disease status, we obtained 69 suspects with $p \leq 0.05$ (Fig 1). Application of an FDR correction of $\leq 20\%$ [23] predicted four *Mtb*-associated metabolites significantly increased in patients with active TB and undetectable in nearly all household contacts ($p < 0.001$ for all). Chemical verification experiments for these metabolites using MS/MS revealed that one was likely a source fragment of the rifampin metabolite 25-deacetyl rifampin. The remaining three significant metabolites were predicted to be phosphatidylglycerol (PG)

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

	Active TB Disease (n = 17)	Household Contacts (n = 16)	P-value ^c
Age [years; median (IQR)] ^a	27 (24–42)	35 (29–48)	0.05
Male sex, n (%)	10 (59%)	6 (38%)	0.22
Tobacco use	13 (76%)	6 (38%)	0.02
High grade AFB sputum smear ^a	1 (6%)	N/A	N/A
>100 Colonies on semi-quantitative culture	7 (41%)	N/A	N/A
Cavitation on chest radiograph	2 (12%)	N/A	N/A
Multidrug-resistant TB ^b	3 (18%)	N/A	N/A

IQR, interquartile range; AFB, acid-fast bacilli; TB, tuberculosis.

a. Sputum smear and culture obtained at study enrollment; high grade smear defined as > 1 AFB per high power field

b. Multidrug resistance was defined as resistance to both isoniazid and rifampin by drug susceptibility testing

c. A Wilcoxon rank sum test was used for comparisons of continuous data and a two-tailed Fisher exact test was used for categorical data

<https://doi.org/10.1371/journal.pone.0205398.t001>

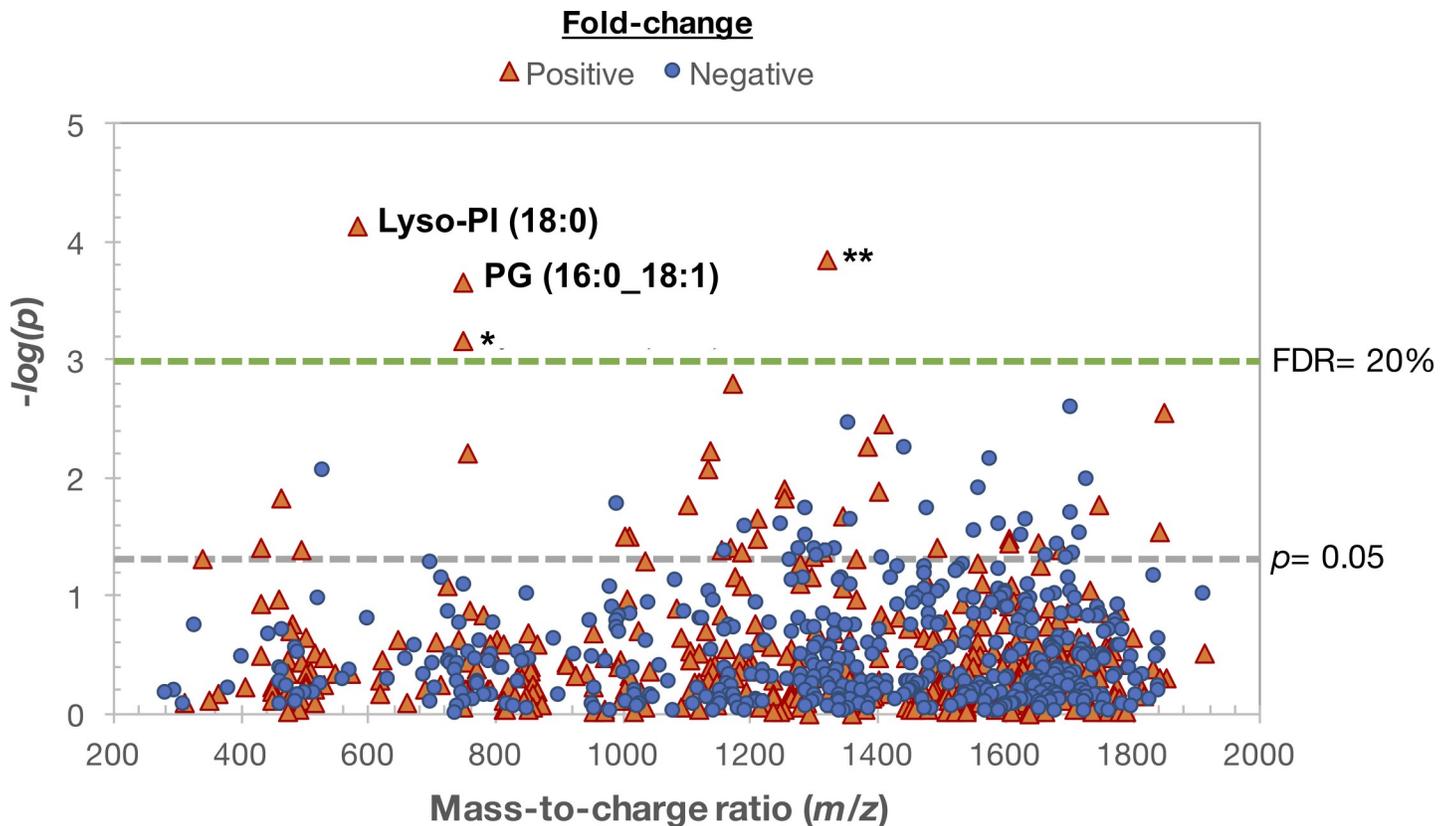


Fig 1. *Mycobacterium tuberculosis* (Mtb) lipid suspects in plasma of adults with active pulmonary TB compared to their household contacts. The Manhattan plot depicts the $-\log p$ statistical analysis of 867 features identified using plasma high-resolution metabolomics (HRM) as *Mtb*-associated metabolite suspects based on accurate mass/charge (m/z) matches. Analysis was done in a cross-sectional comparison of plasma from 17 adults with sputum culture-proven TB disease (within 7 days of diagnosis and initiation of anti-TB drugs) and 16 of their asymptomatic adult household contacts, who were sputum smear and culture negative for *Mtb*. Sixty-nine metabolites were significant at raw $p \leq 0.05$ (points above gray line) and four were significant using a more stringent false discovery rate (FDR) threshold of 20% (red triangles above green line). 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). *Chemical verification experiments revealed the metabolite with m/z 749.3714 was likely a source fragment of 25-desacetyl rifampin. ** m/z 1321.9177 was predicted as acylphosphatidylinositol mannoside (Ac1PIM1) (56:1) based on accurate mass and retention time measurements, but could not be confirmed using tandem MS/MS; PG, phosphatidylglycerol; Lyso-PI, lysophosphatidylinositol.

<https://doi.org/10.1371/journal.pone.0205398.g001>

(16:0_18:1), lysophosphatidylinositol (Lyso-PI) (18:0) and acylphosphatidylinositol mannoside (Ac1PIM1) (56:1) based on accurate m/z and retention time measurements (Fig 2). At least one of these three metabolites was present in plasma of all persons with active TB disease and two or more were detectable in 14 of 17 patients with TB (82%). Among the 16 household contacts, one of the three metabolites was detectable in eight individuals (50%), and only one household contact (6%) had two of these metabolites detectable. Ion fragmentation spectra generated using tandem MS/MS supported the identification of PG (16:0_18:1) and Lyso-PI (18:0) in the plasma of active TB patients (S1 and S2 Figs), while a confirmatory fragmentation spectrum for Ac1PIM1 could not be generated.

TB disease classification accuracy

Using a logistic regression model, the \log_2 intensity value of the three most significant metabolites [PG (16:0_18:1), Lyso-PI (18:0) and m/z 1321.9177] provided excellent classification accuracy of active TB cases. The ROC curve demonstrated an area under the curve (AUC) of 0.97 (95% CI 0.93–1) (Table 2). The AUC was reduced to 0.94 when using only PG (16:0_18:1) and Lyso-PI (18:0) and 0.82 when PG (16:0_18:1) alone (Fig 3).

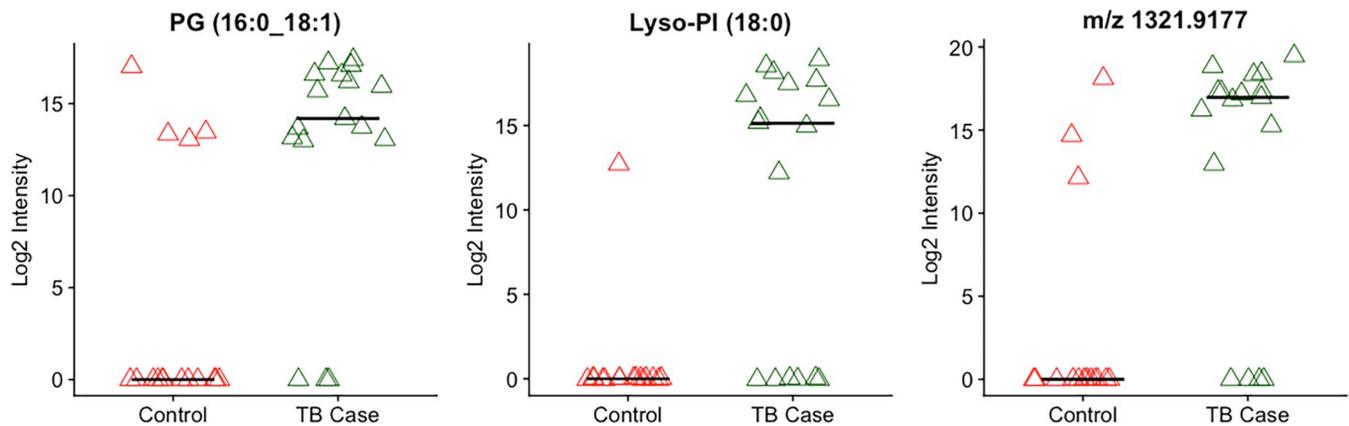


Fig 2. Three most significant *Mtb* lipid matches in plasma. Intensity values of the *Mtb*-associated metabolites phosphatidylglycerol (PG) (16:0_18:1) and lysophosphatidylinositol (Lyso-PI) (18:0), as well as *m/z* 1321.9177 [each $p < 0.001$; active TB cases (green triangles) vs household contacts (red triangles)]. *m/z* 1321.9177 was predicted as acylphosphatidylinositol mannoside (Ac1PIM1) (56:1) based on accurate mass and retention time measurements, but could not be confirmed using tandem MS/MS. Line depicts median values for each subject cohort.

<https://doi.org/10.1371/journal.pone.0205398.g002>

Discussion

Using a suspect screening analysis of *Mtb*-associated metabolite features, we found that metabolites matching the mass and expected retention times of PG (16:0_18:1), Lyso-PI (18:0), and Ac1PIM1 (56:1) were the most significantly increased in the plasma of patients with active pulmonary TB and were largely absent from their asymptomatic household contacts without active TB. The identities of the PG (16:0_18:1) and Lyso-PI (18:0) were supported by tandem MS/MS analysis. Alone or together, these three metabolites demonstrated excellent classification accuracy for active TB cases (AUC 0.97) and therefore hold promise for further development as a metabolomics signature of active TB.

These results also demonstrate that when using rapidly developing HRM technology to study the metabolome of persons with infectious diseases, focused analysis of low-abundance metabolites putatively derived from the infecting microorganisms may reveal new diagnostic targets. In our previous analysis, statistical significance was determined by applying a stringent FDR threshold of 5% to all detected metabolites and the PG (16:0_18:1) and Lyso-PI (18:0) molecules described here did not meet that false discovery threshold [7]. However, selecting molecular features by a low p alone assumes all features have an equal pre-test probability of being biologically meaningful in the context of the disease process studied [24]. Molecular associations in a select cohort of patients may have little to do with disease pathogenesis and these associations are often difficult to replicate in larger studies [25]. Therefore, integration of

Table 2. Area under the receiver operator characteristic curve for varying combinations of *Mycobacterium tuberculosis* lipid suspects.

Metabolites	Area Under the Curve	95% Confidence Interval
PG (16:0_18:1)	0.82	0.68–0.97
PG (16:0_18:1), Lyso-PI (18:0)	0.94	0.88–1
PG (16:0_18:1), Lyso-PI (18:0), <i>m/z</i> 1321.9177*	0.97	0.93–1

* *m/z* 1321.9177 was predicted as acylphosphatidylinositol mannoside (Ac1PIM1) (56:1) based on accurate mass and retention time measurements, but could not be confirmed using tandem MS/MS; PG, phosphatidylglycerol; Lyso-PI, lysophosphatidylinositol

<https://doi.org/10.1371/journal.pone.0205398.t002>

ROC Curve For Identification of Active TB Disease

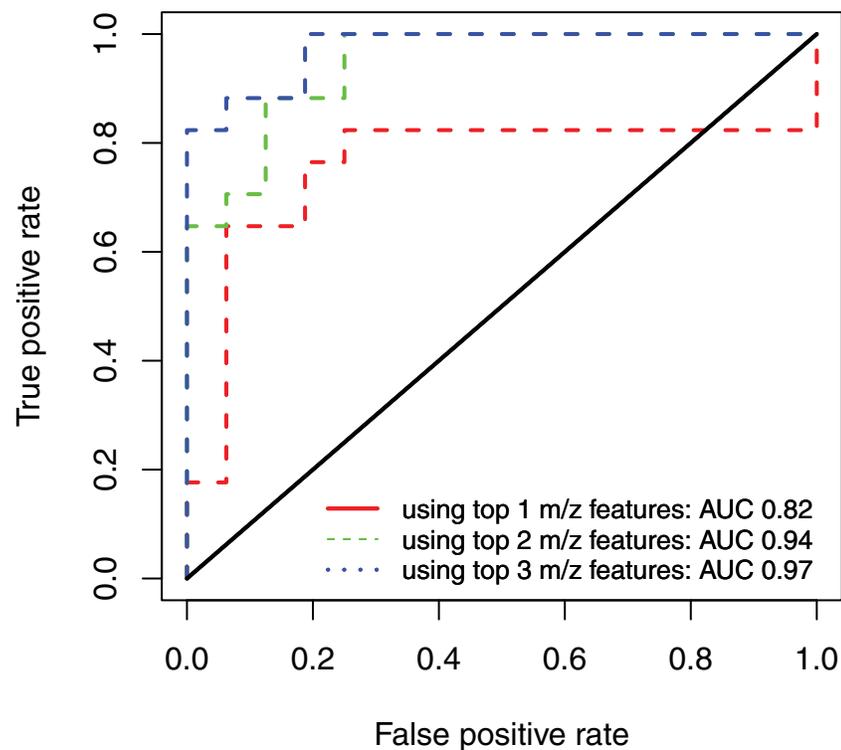


Fig 3. Receiver operating characteristic curve. *M. tuberculosis* lipid suspects phosphatidylglycerol (PG) (16:0_18:1), lysophosphatidylinositol (Lyso-PI) (18:0), and *m/z* 1321.9177 provided excellent classification accuracy for active TB disease [area under the curve (AUC) = 0.97], compared with PG (16:0_18:1) and Lyso-PI (18:0) (AUC = 0.94) or PG (16:0_18:1) alone (AUC = 0.82).

<https://doi.org/10.1371/journal.pone.0205398.g003>

prior information relevant to the disease process of interest provides a strategy to help identify more relevant molecular features [25].

For the present study, we hypothesized that previously published metabolites described as specific products of *Mtb* [16] could be detected in plasma and would distinguish active TB patients from household contacts. This approach led to the discovery of two additional metabolites [PG (16:0_18:1) and Lyso-PI (18:0)] with high predictive value for active TB disease that we excluded in our previous study. These metabolites were previously overlooked when correcting for comparisons of all observed chemical features regardless of potential biologic relevance.

Databases commonly used to annotate metabolites in human studies, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG), the Human Metabolome Database (HMDB) and METLIN, contain few metabolites derived from microorganisms [13–15]. Thus, pathogen-derived molecules may be missed or classified as unknowns. To illustrate, in our previous untargeted metabolomics analysis [7], we found that the *m/z* feature 1321.9177 was significantly higher in patients with active TB disease relative to household controls. However, there are no features matching this *m/z* in the METLIN or LIPD MAPS databases, and the metabolite was initially classified as an unmatched (unknown) metabolite. After using an *Mtb*-specific metabolite library, we found this feature matched the mass and retention time of an acylated

phosphatidylinositol mannoside found in the cell wall of *Mtb* [16]. Although the chemical identity of this feature could not be verified in our current analysis using tandem MS/MS, its association with TB disease raises an interesting possibility that it may be derived from *Mtb* cell wall.

The current study is the first, to our knowledge, to use HRM to comprehensively screen plasma collected from patients with active TB disease for previously identified *Mtb*-associated metabolites. The ultra-high-resolution of the metabolomics methods used has increased sensitivity to detect very low abundance metabolites in biologic samples [26]. The *Mtb* cell wall is constituted by a wide array of glycolipids, making these ideal targets for biomarkers of active TB disease [16, 17, 27, 28]. It is therefore possible that as HRM technology develops, additional metabolites potentially derived from *Mtb* in persons with active TB disease will be identified.

The metabolites differentiating active TB cases from household contacts in the present study may also reveal insights into the host response to *Mtb* infection. PG (16:0_18:1), for example, has been shown to be a strong activator of diverse natural killer T (dNKT) cells during infection with *Mtb* [29]. While PG (16:0_18:1) is found in the cell wall of *Mtb*, it is also found in human pulmonary surfactant. Presentation of both the bacterial regioisomer [PG (18:1/16:0)], as well as the human regioisomer [PG (16:0/18:1)] by CD1d leads to robust pro-inflammatory responses by dNKT cells [29]. Patients with active TB disease may therefore have concomitantly increased plasma concentrations of both endogenously-derived and *Mtb*-derived PG (16:0_18:1) to augment the host immune response. The relative contribution of human and microbial PG (16:0_18:1) to the overall plasma concentration remains unclear and requires further study.

Similarly, Lyso-PI (18:0) is found in both human cells and in *Mtb*; thus, elevated concentrations of this lipid metabolite may derive from the infecting microorganisms and/or the host response to TB disease. Although the abundance of Lyso-PI (18:0) in human cells is low relative to other glycolipids, it has been described as an important cell signaling molecule, acting through G protein-coupled receptor 55 (GPR55) [30]. Lyso-PI (18:0)-induced activation of GPR55 has been shown to activate nuclear factor of activated T-cells (NFAT) and nuclear factor κ of activated B cells (NF- κ B), and has also been associated with activation of wound healing processes in human lung microvascular endothelial cells [31]. Ac1PIM1 is not present in human cells and therefore may be derived from breakdown of the *Mtb* cell wall in the current study. However, this lipid is also found in the cell wall of other types of mycobacteria [27]. Therefore, comparative evaluation of this metabolite in the plasma of individuals during the clinical course of active, symptomatic infection with other mycobacterial species would be of interest.

Although only one of the 16 household contacts in this study had detectable concentrations of two of the three most significant *Mtb* metabolite matches in plasma, half of these individuals had detectable levels of at least one of these metabolites. There are several possible explanations for this finding. Given that household contacts would have been heavily exposed to the active pulmonary TB cases enrolled in the study, some may have had LTBI, which was not tested for in this cohort. There is increasing evidence that persons with LTBI have ongoing organism replication requiring continuing control by the immune system [32, 33], potentially resulting in overlapping plasma metabolic signatures relative to those with active TB. Further investigations are needed to assess the metabolomic profile of individuals with LTBI, targeting the *Mtb*-associated metabolites outlined in this report. Given that PG (16:0_18:1) and lyso-PI (18:0) are also found in human cells and Ac1PIM1 (56:1) is found in the cell wall of other species of mycobacteria [34], future studies should ideally incorporate control groups with non-tuberculous mycobacterial disease and persons with rigorous negative testing for LTBI to further refine the specificity of these metabolites to active TB disease.

This study is subject to several limitations. The sample size for this pilot study was a convenience sample, containing only 17 pulmonary TB cases and 16 household contacts. Due to the small sample size, the three *Mtb*-associated metabolites found to distinguish patients with active TB from their contacts will need to be validated on a larger, prospective cohort of active TB patients and controls with and without LTBI. Plasma HRM studies in humans with active non-tuberculous mycobacterial disease would also be of interest. The current study was conducted at a single center in the country of Georgia with a homogenous (Caucasian) population with a low rate of HIV co-infection. Further work is needed to investigate the presence of these metabolites in other patient populations, including those with HIV co-infection. Additionally, certain *Mtb* lipid metabolites have a larger mass than we queried for identification (i.e. > 2000 m/z) and may have been missed by this analysis; also, our plasma extraction method for LC-MS was not optimized for the most hydrophobic lipid species. Finally, while the identities of PG (16:0_18:1) and Lyso-PI (18:0) were supported by MS/MS confirmation experiments, we were unable to generate interpretable fragmentation spectra for Ac1PIMI (56:1) and this metabolite is therefore reported as a prediction with level five confidence (exact mass of interest) [35]. It is possible the abundance of this metabolite was lower than the threshold necessary to reliably produce interpretable fragmentation spectra.

In future studies, alternative metabolite extraction methods optimized for less polar lipid molecules may increase the yield of *Mtb* lipids in biosamples and allow for additional confirmation of chemical identities. Reproduction of these results in an independent population of control patients both with and without LTBI would also provide additional confidence they are either derived from *Mtb* or related to its pathophysiology. It would also be of interest to analyze patient urine for excretion of these *Mtb* metabolites to better understand the kinetics of detection in relation to TB disease diagnosis and response to therapy. Recent developments using hydrogel nanocage technology suggest the mycobacterial cell wall component lipoarabinomannan can be detected at extremely low concentrations in urine samples [36]. Confirmation of LAM breakdown products in urine, concomitant with LAM, would lend important insight to the utility of this assay as a TB disease biomarker.

In conclusion, we found that among several *Mtb* lipid metabolite m/z matches detected, PG (16:0_18:1) and Lyso-PI (18:0) were the molecules most significantly elevated in the plasma of patients with active TB compared to their asymptomatic household contacts. These metabolites provided excellent classification accuracy for active TB. Our pilot data suggests that plasma HRM may hold significant promise for characterization of *Mtb*-associated metabolites in patients with active TB disease. If confirmed in larger studies that include diverse patient populations and relevant controls, further development of the most discriminatory metabolites may inform development of diagnostic biomarkers for active TB.

Supporting information

S1 File. Full feature table. This file contains the full feature table with metabolite intensities for all samples analyzed in this study organized by mass-to-charge ratio (m/z), retention time, and minimum and maximum m/z .

(ZIP)

S1 Fig. MS/MS fragmentation spectra show positive identification of phosphatidylglycerol (16:0_18:1). MS/MS spectra for m/z 747.7. Presence of fragments 673.3, 281.0 and 255.0 in the plasma sample of an active TB patient indicate a match to the [M-H] adduct for phosphatidylglycerol (PG) (16:0_18:1).

(TIFF)

S2 Fig. MS/MS fragmentation spectra show positive identification of lysophosphatidylinositol (18:0). MS/MS spectra for m/z 599.5. Presence of fragments 315.0, 283.3 and 152.8 in the plasma sample of an active TB patient indicate a match to the [M-H] adduct for lysophosphatidylinositol (Lyso-PI) (18:0). (TIFF)

Author Contributions

Conceptualization: Jeffrey M. Collins, Douglas I. Walker, Dean P. Jones, Nestani Tukvadze, Russell R. Kempker, Henry M. Blumberg, Thomas R. Ziegler.

Data curation: Jeffrey M. Collins, Douglas I. Walker, ViLinh T. Tran, Karan Uppal, Jennifer K. Frediani, Kirk A. Easley, Neeta Shenvi.

Formal analysis: Jeffrey M. Collins, Douglas I. Walker, Ken H. Liu, Karan Uppal, Jennifer K. Frediani, Kirk A. Easley, Neeta Shenvi, Thomas R. Ziegler.

Funding acquisition: Dean P. Jones, Nestani Tukvadze, Henry M. Blumberg, Thomas R. Ziegler.

Investigation: Jeffrey M. Collins, Douglas I. Walker, Dean P. Jones, Nestani Tukvadze, Ken H. Liu, ViLinh T. Tran, Jennifer K. Frediani, Manoj Khadka, Eric A. Ortlund, Russell R. Kempker, Henry M. Blumberg, Thomas R. Ziegler.

Methodology: Douglas I. Walker, Ken H. Liu, ViLinh T. Tran, Karan Uppal, Jennifer K. Frediani, Manoj Khadka, Eric A. Ortlund, Russell R. Kempker, Thomas R. Ziegler.

Project administration: Dean P. Jones, Nestani Tukvadze, Jennifer K. Frediani, Kirk A. Easley, Thomas R. Ziegler.

Resources: Douglas I. Walker, Dean P. Jones, Nestani Tukvadze, Henry M. Blumberg, Thomas R. Ziegler.

Software: Karan Uppal.

Supervision: Dean P. Jones, Russell R. Kempker, Henry M. Blumberg, Thomas R. Ziegler.

Visualization: Jeffrey M. Collins, Karan Uppal, Manoj Khadka, Eric A. Ortlund.

Writing – original draft: Jeffrey M. Collins.

Writing – review & editing: Jeffrey M. Collins, Douglas I. Walker, Dean P. Jones, Nestani Tukvadze, Ken H. Liu, ViLinh T. Tran, Karan Uppal, Jennifer K. Frediani, Kirk A. Easley, Neeta Shenvi, Manoj Khadka, Eric A. Ortlund, Russell R. Kempker, Henry M. Blumberg, Thomas R. Ziegler.

References

1. World Health Organization. Global tuberculosis report 2016. 2017. Available from: http://www.who.int/tb/publications/global_report/en.
2. Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, Cunningham J, et al. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis.* 2006; 6: 664–674. [https://doi.org/10.1016/S1473-3099\(06\)70602-8](https://doi.org/10.1016/S1473-3099(06)70602-8) PMID: 17008175
3. Pai M, Schito M. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. *J Infect Dis.* 2015; 211 Suppl 2: S21–28.
4. Wallis RS, Kim P, Cole S, Hanna D, Andrade BB, Maeurer M, et al. Tuberculosis biomarkers discovery: developments, needs, and challenges. *Lancet Infect Dis.* 2013; 13: 362–372. [https://doi.org/10.1016/S1473-3099\(13\)70034-3](https://doi.org/10.1016/S1473-3099(13)70034-3) PMID: 23531389

5. Lienhardt C, Lonnroth K, Menzies D, Balasegaram M, Chakaya J, Cobelens F, et al. Translational research for tuberculosis elimination: priorities, challenges, and actions. *PLoS Med.* 2016; 13: e1001965. <https://doi.org/10.1371/journal.pmed.1001965> PMID: 26933883
6. Go YM, Walker DI, Liang Y, Uppal K, Soltow QA, Tran V, et al. Reference standardization for mass spectrometry and high-resolution metabolomics applications to exposome research. *Toxicol Sci.* 2015; 148: 531–542. <https://doi.org/10.1093/toxsci/kfv198> PMID: 26358001
7. Frediani JK, Jones DP, Tukvadze N, Uppal K, Sanikidze E, Kipiani M, et al. Plasma metabolomics in human pulmonary tuberculosis disease: a pilot study. *PloS One.* 2014; 9: e108854. <https://doi.org/10.1371/journal.pone.0108854> PMID: 25329995
8. Collins JM, Kempker RR, Ziegler TR, Blumberg HM, Jones DP. Metabolomics and mycobacterial disease: don't forget the bioinformatics. *Ann Am Thorac Soc.* 2016; 13: 141–142. <https://doi.org/10.1513/AnnalsATS.201510-676LE> PMID: 26730872
9. Jones DP. Sequencing the exposome: A call to action. *Toxicol Rep.* 2016; 3: 29–45. <https://doi.org/10.1016/j.toxrep.2015.11.009> PMID: 26722641
10. Uppal K, Walker DI, Liu K, Li S, Go YM, Jones DP. Computational metabolomics: a framework for the million metabolome. *Chem Res Toxicol.* 2016; 29: 1956–1975. <https://doi.org/10.1021/acs.chemrestox.6b00179> PMID: 27629808
11. Lau SK, Lee KC, Curreem SO, Chow WN, To KK, Hung IF, 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; 53: 3750–3759. <https://doi.org/10.1128/JCM.01568-15> PMID: 26378277
12. Mahapatra S, Hess AM, Johnson JL, Eisenach KD, DeGroot MA, Gitta P, et al. A metabolic biosignature of early response to anti-tuberculosis treatment. *BMC Infect Dis.* 2014; 14: 53. <https://doi.org/10.1186/1471-2334-14-53> PMID: 24484441
13. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, et al. METLIN: a metabolite mass spectral database. *Ther Drug Monit.* 2006; 27: 747–751.
14. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 2007; 35 Database issue: D521–526. <https://doi.org/10.1093/nar/gkl923> PMID: 17202168
15. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 1999; 27: 29–34. PMID: 9847135
16. Sartain MJ, Dick DL, Rithner CD, Crick DC, Belisle JT. Lipidomic analyses of *Mycobacterium tuberculosis* based on accurate mass measurements and the novel "Mtb LipidDB". *J Lipid Res.* 2011; 52: 861–872. <https://doi.org/10.1194/jlr.M010363> PMID: 21285232
17. Layre E, Sweet L, Hong S, Madigan CA, Desjardins D, Young DC, et al. A comparative lipidomics platform for chemotaxonomic analysis of *Mycobacterium tuberculosis*. *Chem Biol.* 2011; 18: 1537–1549. <https://doi.org/10.1016/j.chembiol.2011.10.013> PMID: 22195556
18. Tukvadze N, Sanikidze E, Kipiani M, Hebbar G, Easley KA, Shenvi N, et al. High-dose vitamin D3 in adults with pulmonary tuberculosis: a double-blind randomized controlled trial. *Am J Clin Nutr.* 2015; 102: 1059–1069. <https://doi.org/10.3945/ajcn.115.113886> PMID: 26399865
19. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics.* 2013; 9 Suppl 1: S132–143.
20. Yu T, Park Y, Johnson JM, Jones DP. apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics.* 2009; 25: 1930–1936. <https://doi.org/10.1093/bioinformatics/btp291> PMID: 19414529
21. Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, Yu T, et al. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC bioinformatics.* 2013; 14: 15. <https://doi.org/10.1186/1471-2105-14-15> PMID: 23323971
22. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015; 43: e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
23. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med.* 1990; 9: 811–818. PMID: 2218183
24. Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics.* 2006; 2: 171–196.
25. Lee SI, Celik S, Logsdon BA, Lundberg SM, Martins TJ, Oehler VG, et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat Commun.* 2018; 9: 42. <https://doi.org/10.1038/s41467-017-02465-5> PMID: 29298978
26. Haas CT, Roe JK, Pollara G, Mehta M, Noursadeghi M. Diagnostic 'omics' for active tuberculosis. *BMC Med.* 2016; 14: 37. <https://doi.org/10.1186/s12916-016-0583-9> PMID: 27005907

27. Mishra AK, Driessen NN, Appelmek BJ, Besra GS. Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in Mycobacterium tuberculosis physiology and host-pathogen interaction. *FEMS Microbiol Rev.* 2011; 35: 1126–1157. <https://doi.org/10.1111/j.1574-6976.2011.00276.x> PMID: 21521247
28. Shui G, Bendt AK, Jappar IA, Lim HM, Laneelle M, Herve M, 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. <https://doi.org/10.1002/emmm.201100185> PMID: 22147526
29. Tatituri RV, Watts GF, Bhowruth V, Barton N, Rothchild A, Hsu FF, et al. Recognition of microbial and mammalian phospholipid antigens by NKT cells with diverse TCRs. *Proc Natl Acad Sci U S A.* 2013; 110: 1827–32. <https://doi.org/10.1073/pnas.1220601110> PMID: 23307809
30. Oka S, Nakajima K, Yamashita A, Kishimoto S, Sugiura T. Identification of GPR55 as a lysophosphatidylinositol receptor. *Biochem Biophys Res Commun.* 2007; 362: 928–934. <https://doi.org/10.1016/j.bbrc.2007.08.078> PMID: 17765871
31. Kargl J, Brown AJ, Andersen L, Dorn G, Schicho R, Waldhoer M, et al. A selective antagonist reveals a potential role of G protein-coupled receptor 55 in platelet and endothelial cell function. *J Pharmacol Exp Ther.* 2013; 346: 54–66. <https://doi.org/10.1124/jpet.113.204180> PMID: 23639801
32. Esmail H, Lai RP, Lesosky M, Wilkinson KA, Graham CM, Coussens AK, et al. Characterization of progressive HIV-associated tuberculosis using 2-deoxy-2-[18F]fluoro-D-glucose positron emission and computed tomography. *Nat Med.* 2016; 22: 1090–1093. <https://doi.org/10.1038/nm.4161> PMID: 27595321
33. Ghesani N, Patrawalla A, Lardizabal A, Salgame P, Fennelly KP. Increased cellular activity in thoracic lymph nodes in early human latent tuberculosis infection. *Am J Respir Crit Care Med.* 2014; 189: 748–750. <https://doi.org/10.1164/rccm.201311-1976LE> PMID: 24628316
34. Crellin PK, Luo CY, Morita YS. Metabolism of plasma membrane lipids in mycobacteria and corynebacteria. In: Baez RV, editor. *Lipid Metabolism.* In Tech; 2013. Available from <https://www.intechopen.com/books/lipid-metabolism/metabolism-of-plasma-membrane-lipids-in-mycobacteria-and-corynebacteria>.
35. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol.* 2014; 48: 2097–2098. <https://doi.org/10.1021/es5002105> PMID: 24476540
36. Paris L, Magni R, Zaidi F, Araujo R, Saini N, Harpole M, et al. Urine lipoarabinomannan glycan in HIV-negative patients with pulmonary tuberculosis correlates with disease severity. *Sci Transl Med.* 2017; 9 (420).