

Sputum proteomic analysis for distinguishing between pulmonary tuberculosis and non-tuberculosis using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS): preliminary results

Thi Loi Dao, van Thuan Hoang, Tran Duc Anh Ly, Jean Christophe Lagier, Sophie Alexandra Baron, Didier Raoult, Philippe Parola, Johan Courjon, Pierre Marty, Hervé Chaudet, et al.

▶ To cite this version:

Thi Loi Dao, van Thuan Hoang, Tran Duc Anh Ly, Jean Christophe Lagier, Sophie Alexandra Baron, et al.. Sputum proteomic analysis for distinguishing between pulmonary tuberculosis and non-tuberculosis using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS): preliminary results. Clinical Microbiology and Infection, 2021, 27 (11), pp.1694.e1-1694.e6. 10.1016/j.cmi.2021.02.031. hal-03636901

HAL Id: hal-03636901 https://amu.hal.science/hal-03636901

Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1	Sputum proteomic analysis for distinguishing between pulmonary tuberculosis and non-
2	tuberculosis using matrix-assisted laser desorption ionization time-of-flight mass
3	spectrometry (MALDI-TOF MS): preliminary results
4	Running title: MALDI-TOF MS using for distinguishing PTB and non-TB
5	Thi Loi Dao ^{1, 2,3} , Van Thuan Hoang ^{1, 2,3} , Tran Duc Anh Ly ^{1, 2} , Jean Christophe Lagier ^{2, 4} , Sophie
6	Alexandra Baron ^{2, 4} , Didier Raoult ^{2, 4} , Philippe Parola ^{1, 2} , Johan Courjon ⁵ , Pierre Marty ⁶ , Hervé
7	Chaudet ^{1, 2} , Philippe Gautret ^{1, 2} *
8	¹ Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, Marseille, France
9	² IHU-Méditerranée Infection, Marseille, France
10	³ Thai Binh University of Medicine and Pharmacy, Thai Binh, Vietnam
11	⁴ Aix Marseille Univ, IRD, AP-HM, MEPHI, Marseille, France
12	⁵ Université Côte d'Azur, U1065, Centre Méditerranéen de Médecine Moléculaire, C3M,
13	Virulence Microbienne et Signalisation Inflammatoire, Inserm, Nice, France; Infectiologie,
14	Hôpital de l'Archet, Centre Hospitalier Universitaire de Nice, Nice, France
15	⁶ Université Côte d'Azur, Inserm, C3M, Nice Cedex 3, France; Parasitologie-Mycologie, Centre
16	Hospitalier Universitaire l'Archet, Nice Cedex 3, France.
17	*Corresponding author:
18	Philippe Gautret
19	VITROME, Institut Hospitalo-Universitaire Méditerranée Infection, 19-21 Boulevard Jean
20	Moulin 13385 Marseille Cedex 05, France. Phone: + 33 (0) 4 13 73 24 01. Fax: + 33 (0) 4 13 73

21 24 02. E-mail address: philippe.gautret@club-internet.fr

Abstract (246/250 words)

Objectives: To evaluate the feasibility and diagnostic contribution of protein profiling using matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) applied to sputum to diagnose pulmonary tuberculosis.

Methods: Sputum samples collected from patients suspected of having pulmonary tuberculosis were analysed using MALDI-TOF MS. Using the differentially expressed protein peaks, we compared three groups of patients, including those with confirmed pulmonary tuberculosis (PTB), those without tuberculosis but with a lower respiratory tract infection (non-TB LRTI), and those without tuberculosis and without an LRTI (non-TB controls).

Results: A total of 102 patients included 35 PTB, 36 non-TB LRTI and 31 non-TB controls. The model differentiated between the PTB patients and the non-TB controls using the 25 most differentially expressed protein peaks, with a sensitivity of 97%, 95%CI=[85%–100%] and a specificity of 77%, 95%CI=[59%-90%]. The model distinguished the PTB patients from the non-TB LRTI patients using the 10 most differentially expressed protein peaks, with a sensitivity of 80%, 95%CI=[63%–92%] and a specificity of 89%, 95%CI=[74%–97%]. We observed that the negative predictive value of MALDI-TOF MS sputum analysis was higher (96%, 95%CI=[80%–100%]) than that of direct sputum microscopic examination and sputum culture (78%, 95%CI=[62%–89%]) for non-TB controls. When MALDI-TOF MS sputum analysis and direct microscopic examination were combined, the negative predictive value reached 94%, 95%CI=[80%–99%] for non-TB LRTI patients.

Conclusions: These results suggest that MALDI-TOF MS sputum analysis coupled with microscopic examination could be used as a screening tool for diagnosing pulmonary TB.

Keywords: MALDI-TOF; proteomic; tuberculosis; respiratory infection; sensitivity; specificity

Text (2343/2500 words)

Introduction

In 2018, the World Health Organization reported 10 million new cases of tuberculosis and 1.2 million deaths around the world, making the disease a significant public health concern [1]. The diagnosis of pulmonary tuberculosis is currently based on the detection of *M. tuberculosis* complex by direct examination with microscopy or by detection of specific DNA sequences by PCR followed by isolation of the mycobacteria by culture.

Sputum contains a wide variety of proteins the composition of which most closely reflects the condition of the lungs. Another advantage is that sputum is easy to sample in a non-invasive way [2]. Numerous studies have analysed the protein profile of sputum to identify protein markers as a new alternative diagnostic method [3, 4]. The vast majority of proteomic studies of sputum have been conducted in patients with inflammatory diseases including asthma and chronic obstructive pulmonary disease (COPD) but rarely in patients with lung infections [5]. Several authors suggested that specific protein profiles could represent the signature of pulmonary tuberculosis [6-8].

The objective of this study was to evaluate the feasibility and diagnostic contribution of protein profiling using MALDI-TOF MS when directly applied to sputum from patients with confirmed pulmonary tuberculosis, patients with lower respiratory tract infections other than tuberculosis, and patients with no lower respiratory tract infection in whom tuberculosis had been excluded (controls). We hypothesised that rapid routine and blind MALDI-TOF MS analysis could accurately distinguish these three populations.

Methods

Study population

The population was recruited between January 2019 and June 2020 among hospitalized patients at the IHU, and who were suspected of having pulmonary tuberculosis at the time of admission. In line with the WHO recommendations for screening for tuberculosis [9], the population included subjects with clinical symptoms suggestive of pulmonary tuberculosis, as well as subjects with an epidemiological context putting them at risk of pulmonary tuberculosis, people living with HIV, and people who were immunosuppressed.

Based on the final diagnosis, patients were divided into three groups: (1) the pulmonary tuberculosis group (PTB), in which pulmonary tuberculosis had been confirmed by direct microscopic examination and/or Gene Xpert[®] MTB-RIF and/or culture; (2) the lower respiratory tract infections other than tuberculosis (non-TB-LRTI) group, including patients with an LRTI whose tuberculosis tests had all been negative and who received a final diagnosis of pneumonia or exacerbation of COPD with a secondary infection; and (3) non-tuberculosis, non-LRTI (non-TB control) group including patients whose tuberculosis tests were all negative and had excluded any LRTIs. A flow diagram of the study is presented in Supplementary Figure 1. Details of these three groups are shown in Supplementary Table S1.

Clinical samples

Sputum samples were collected as part of the routine diagnostic activity for pulmonary tuberculosis at our clinical microbiology laboratory (IHU). Sputum specimens were collected in accordance with the relevant guidelines and regulations [10]. Microscopic detection of acid-fast

bacilli (AFB) was performed by laboratory technicians using a commercially available kit featuring Kinyoun staining (RAL Diagnostics, Toulouse, France). Sputum specimens were inoculated onto C-Top COSMO medium (Eurobio SA, Les Ulis, France) after decontamination. GeneXpert[®] MTB/RIF (Cepheid, Sunnyvale, CA, United-States) was used to detect tuberculosis mycobacteria and rifampicin resistance genes in sputum, according to the manufacturer's recommendations.

The remaining sputum material was kept into a dry, sterile container and stored at -80°C until processing for mass spectrometry.

Mass spectrometry

For each sample, a volume of 2.0 μ L of sputum was directly deposited on four wells (0.5 μ L each well) of a MALDI 96 steel target, then dried at room temperature. Deposits were next coated with 1.0 μ L of CHCA matrix solution (α -cyano-4-hydroxycinnamic acid diluted into 500 μ L of acetonitrile, 250 μ L of HPLC grade water and 250 μ L of 10% trifluoroacetic acid and). The plate was dried at room temperature, then was placed into a Microflex LT MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). Each sputum sample generated four spectra issued from four deposits. Spectra were recorded within a mass range from 2 to 20 kDa. Each spectrum was performed in four regions of the same spot and next automatically acquired using the AutoXecute acquisition control in FlexControl software 3.0 (Bruker Daltonics). The BioTyper-RTC v.3.0 software (Bruker Daltonics GmbH) was used to import all mass spectra. The calibration of the mass spectrometer was fully automatic and executed with a commercial solution (BTS: Bacterial Test Standard, Bruker Daltonics GmbH, Germany) and the strategy was completely automated (BioTyper-RTC). The quality criteria of the spectra were based on: (1) -

horizontal baseline curve, (2) - presence of visually identifiable peaks and (3) - intensity above 10^4 arbitrary units (AU).

Bioinformatics analysis

All spectra were analysed using R version 4.0.2 [11, 12] (Supplementary data). Differences between the two analysed groups were evalued on the discriminant peak identification list using binary predictors. We executed a discriminant analysis between groups using the Binary Discriminant Analysis method [13] for create the list of discriminant peaks for each specimen. The sensitivity (Se) and specificity (Sp) were calculated by comparing the results of GeneXpert on the same sputum specimens used in the reference method.

Ethics

Mass spectrometry analysis was performed on sputum samples that were collected in the context of routine diagnosis. No additional samples were collected for this study. Clinical data were retrospectively retrieved from medical files and anonymised before analysis. Ethical approval was obtained from the Marseille Institutional Review Board and Ethics Committee (No. 2018-015).

Results

Characteristics of the population

A total of 102 patients were included. Thirty-five cases with confirmed PTB, 36 non-TB LRTI, and 31 non-TB controls were enrolled. Patient details are listed in Table 1 and Supplementary Table S1. PTB and non-TB controls were middle-aged patients, while non-TB LRTI patients

were significantly older. The prevalence of clinical symptoms presented by patients varied significantly between groups. In the non-TB control group, 16/31 patients presented clinical symptoms. Most (28/36) patients with non-TB LRTI had acute community-acquired pneumonia.

Differentially expressed peptide peaks

All samples were analysed using MALDI-TOF-MS, and protein profiles were obtained from each sample in the range of 2–20 kDa. Three representative spectra of three groups are shown in Supplementary Figure 2. To distinguish sputum samples collected from PTB patients from the non-TB controls, a total of 86 protein peaks were retrieved from the sputum of 35 PTB patients and 31 non-TB controls. Of these 86 peaks, 11 were significantly different between the two groups (p< 0.05) (Supplementary Table S2).

To distinguish PTB from the non-TB LRTI sputum samples, a total of 92 protein peaks were retrieved from the sputum of 35 PTB patients and 36 non-TB LRTI patients. Of these 92 peaks, eight were significantly different between the two groups (p < 0.05) (Supplementary Table S3).

Protein peaks selected for the decision model of classification

Using the 25 most differentially expressed protein peaks, it was observed that the model distinguished between the PTB patients and the non-TB controls with a sensitivity of 97%, 95% confidence interval (CI) =[85%–100%] and a specificity of 77%, 95%CI=[59%–90%], (Youden index = 0.75) (Table 2 and Supplementary Figure 3). Of the top 25 ranking peaks, 15 (m/z values: 4161 Da, 2065 Da, 4137 Da, 6892 Da, 5155 Da, 2128 Da, 2202 Da, 2110 Da, 4355 Da, 6188 Da, 2297 Da, 5227 Da, 5002 Da, 12697 Da and 4283 Da) had higher levels of intensities in the PTB group, and 10 (m/z values: 4901 Da, 2316 Da, 14688 Da, 11769 Da, 11736 Da, 2280 Da, 7348 Da, 3629 Da, 2786 Da and 3671 Da) had higher levels of intensity in the non-TB

controls group (Figure 1). A single PTB patient (PTB22) did not clusterise with other PTB patients. This 21-year-old male patient originating from Senegal had a final diagnosis of pulmonary and pleural tuberculosis. This patient was negative for AFB under direct microscopic examination, but positive for *M. tuberculosis* using GeneXpert[®] and sputum culture.

By using the 10 most differentially expressed protein peaks (Figure 2), we found that the model distinguished the PTB patients from the non-TB LRTI patients with a sensitivity of 80% [95%CI=63%–92%] and a specificity of 89% [95%CI=74%–97%], (Youden index= 0.69) (Table 3 and Supplementary Figure 4). Of the top 10 ranking peaks, seven (m/z values: 5687 Da, 5728 Da, 6821 Da, 4242 Da, 6948 Da, 3930 Da, and 5531 Da) had higher levels of intensity in the non-TB LRTI group, and three (m/z values: 4697 Da, 4160 and 4137 Da) had higher levels of intensity in the PTB group (Figure 2).

Evaluation of different diagnostic methods

The performance of the decision model using MALDI-TOF MS, staining methods and culture in comparison to GeneXpert[®] is presented in Supplementary Tables 4 and 5, along with their sensitivity, specificity, negative predictive value (NPV) values and positive predictive value (PPV). For PTB and non–TB control cases, the sensitivity /specificity was 97% (95%CI=[85%–100%])/77% (95%CI=[59%–90%]) for MALDI-TOF-MS using 25 peaks or when MALDI-TOF-MS and microscopic examination of sputum was combined, 74% (95%CI=[57%–88%])/100% (95%CI=[89%–100%]) for culturing and 74% (95%CI=[57%–88%])/100% (95%CI=[89%–100%]) for microscopic examination of sputum. The PPV was 83% (95%CI=[68%–93%]), 100% (95%CI=[87%–100%]) and 100% (95%CI=[87%–100%]), respectively, and the NPV was 96% (95%CI=[80%–100%]), 78% (95%CI=[62%–89%]) and 78% (95%CI=[62%–89%]), respectively.

For PTB and non-TB LRTI cases, the sensitivity /specificity was 80% (95%CI=[63%–92%])/89% (95%CI=[74%–97%]) for MALDI-TOF MS using 10 peaks, 74% (95%CI=[57%–88%])/100% (95%CI=[90%–100%]) for culturing and 74% (95%CI=[57%–88%])/100% (95%CI=[90%–100%]) for microscopic examination of sputum. The PPV was 88% (95%CI=[71%–96%]), 100% (95%CI=[87%–100%]) and 100% (95%CI=[87%–100%]), respectively, and the NPV was 82% (95%CI=[66%–92%]), 80% (95%CI=[65%–90%]) and 80% (95%CI=[65%–90%]), respectively. When microscopic examination and MALDI-TOF MS were combined, the sensitivity/ specificity was 94% (95%CI=[81%–99%])/89% (95%CI=[74%–97%]) and the PPV /NPV were 89% (95%CI=[75%–97%])/94% (95%CI=[80%–99%]).

By using the same method for classification, we compared the two non-TB groups with 25 most differentially expressed protein peaks. The results are shown in Supplementary Figure 5.

Discussion

Of the 102 patients included in this study (35 with confirmed PTB, 31 non-TB controls, and 36 non-TB LRTI), based on the 25 and 10 most differentially expressed protein peaks, we found that the models distinguished PTB patients from non-TB controls and from non-TB LRTI patients with a sensitivity of 97% and 80% and a specificity of 77% and 89%, respectively. The signal noise ratio threshold 2 was used allowing considering exhaustive peaks and increasing the ability to select peaks to analyze, as previously studied [14].

In the past, other studies [6-8] used more resolving methods and could not be applied to a routine diagnostic test. In our work, we did not study for proteomic biomarkers identification. We hypothesized that rapid routine and blinded MALDI-TOF analysis could precisely classify these

three groups of patients. We chose a method following as narrowly as possible the standard MALDI-TOF protocol that is used in routine diagnostic.

There are few studies on the use of MALDI-TOF MS directly from biological samples as a diagnostic tool in microbiology. One report showing that direct profiling of small proteins in cerebrospinal fluid (CSF) samples by MALDI-TOF MS may enable the diagnosis of enterovirus (EV) meningitis in comparison to positive and negative CSF samples for EV [15].

We observed that the NPV of sputum MALDI-TOF MS analysis alone was higher (96%) than that of direct sputum microscopic examination and sputum culture for non-TB controls (78%). The NPV of sputum MALDI-TOF MS analysis was only 82% for non-TB LRTI patients (compared to 80% for direct sputum microscopic examination and sputum culture). However, when sputum MALDI-TOF MS analysis and direct microscopic examination were combined, the NPV reached 94%. This result suggests that sputum MALDI-TOF MS analysis coupled with microscopic examination could be used as a screening tool for the diagnosis of PTB.

Our study has several limitations, including its small sample size and the fact that it was conducted on frozen samples. The total number of most differentially expressed peaks (10 and 25) may be considered as somewhat high, taking in account the number of patients in each group, but it should be underlined that our method considered both peaks specifically present or absent in any group. Also, we did not identify the specific biomarkers responsible for each peak or group of peaks. There were significant differences of age in the three groups of patients that may have represented a confounding factor. Nevertheless, our study was performed using non-invasive clinical samples from patients suspected of having pulmonary tuberculosis, did not require complex technical expertise and could be conducted by any laboratory personnel. In developing countries, the use of MALDI-TOF MS is limited due to the high price. Yet, when

purchased as part of technology platforms with multiple uses in bacteriology, mycology and even entomology, the running costs are much less expensive including the cost of consumables and labour [16, 17].

Future studies on fresh sputum samples would be necessary to carry out to assess the optimal conditions of samples for routine diagnostic tests. Studies conducted on a larger sample of patients with confirmed PTB are needed to confirm our results. It would be of interest to include patients with non-tuberculosis *Mycobacterium* infections as control groups. Finally, there is a need to refine the performance of the method with more than one specimen from a patient at a given time and samples obtained during follow-up in treated patients.

Our study is based on very preliminary results and further investigation must be conducted to confirm the general principle and experimental improvements are needed to strengthen the method. Nevertheless we obtained encouraging first results suggesting that analysis of sputum protein profiles using MALDI-TOF MS directly on samples may enable PTB patients to be distinguished from non-TB patients with or without an LRTI with a good NPV. Because the rapid diagnosis of PTB remains a challenge, the further development of a simple-to-use, rapid and cost-effective tool based on the use of MALDI-TOF MS directly on the sputum would be an attractive screening test, in combination with direct microscopy, in central laboratory settings in future, if results are confirmed.

Authors' Contributions

Conceptualization: Philippe Gautret, Didier Raoult, Pierre Marty

Methodology: Thi Loi Dao, Hervé Chaudet, Philippe Gautret

Laboratory technic: Thi Loi Dao, Van Thuan Hoang, Tran Duc Anh Ly

Data collection: Thi Loi Dao, Jean-Christophe Lagier, Philippe Parola, Johan Courjon

Formal analysis and investigation: Thi Loi Dao, Hervé Chaudet

Writing original draft and preparation: Thi Loi Dao, Philippe Gautret

Writing – review and editing: Thi-Loi Dao, Tran Duc Anh Ly, Jean Christophe Lagier, Sophie Alexandra Baron, Didier Raoult, Philippe Parola, Johan Courjon, Pierre Marty, Hervé Chaudet

Supervision: Philippe Gautret

Funding

This study was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection, the French National Research Agency under the "Investissements d'avenir" programme, reference ANR-10-IAHU-03, the Région Provence Alpes Côte d'Azur and European ERDF PRIMI funding.

Acknowledgments

We are grateful to Michel Drancourt for critically reviewing the manuscript.

Conflict of Interest

The authors declare that they have no conflicts of interest and that they did not receive support by the industry in the last three years.

Availability of data and materials

The datasets generated and/or analysed during the current study are available from the corresponding author [P.G.] on reasonable request.

References

[1]. World Health Organization. Global tuberculosis report 2019.https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1.

[2]. Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC et al. Sputum proteomics in inflammatory and suppurative respiratory diseases. Am J Respir Crit Care Med. 2008; 178:444–52.

[3]. Paone G, Leone V, Conti V, De Marchis L, Ialleni E, Graziani C et al. Blood and sputum biomarkers in COPD and asthma: a review. Eur Rev Med Pharmacol Sci. 2016; 20:698–708.

[4]. Dong T, Santos S, Yang Z, Yang S, Kirkhus NE. Sputum and salivary protein biomarkers and point-of-care biosensors for the management of COPD. Analyst. 2020; 145:1583–604.

[5]. Wheelock CE, Goss VM, Balgoma D, Nicholas B, Brandsma J, Skipp PJ et al. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. Eur Respir J. 2013; 42:802–25.

[6]. Fu YR, Yi ZJ, Guan SZ, Zhang SY, Li M. Proteomic analysis of sputum in patients with active pulmonary tuberculosis. Clin Microbiol Infect. 2012; 18:1241–47.

[7]. Bishwal SC, Das MK, Badireddy VK, Dabral D, Das A, Mahapatra AR et al. Sputum Proteomics Reveals a Shift in Vitamin D-binding Protein and Antimicrobial Protein Axis in Tuberculosis Patients. Sci Rep. 2019; 9:1036.

[8]. Mateos J, Estévez O, González-Fernández Á, Anibarro L, Pallarés Á, Reljic R et al. Highresolution quantitative proteomics applied to the study of the specific protein signature in the sputum and saliva of active tuberculosis patients and their infected and uninfected contacts. J Proteomics. 2019; 195:41–52.

[9].WorldHealthOrganization.2015.https://apps.who.int/iris/bitstream/handle/10665/181164/9789241549172_eng.pdf?sequence=1

[10]. El Khéchine A, Drancourt M. Diagnosis of pulmonary tuberculosis in a microbiological laboratory. Med Mal Infect. 2011; 41:509–17.

[11]. R Core Team. R foundation for Statistical Computing, Vienna, Austria, 2020. Available from: https://www.R-project.org/

[12]. Gibb S, Strimmer K. MALDIquant: a versatile R package for the analysis of mass spectrometry data. Bioinformatics. 2012; 28:2270–1.

[13]. Gibb S, Strimmer K. Differential protein expression and peak selection in mass spectrometry data by binary discriminant analysis. Bioinformatics. 2015; 31:3156–62.

[14]. Giraud-Gatineau A, Texier G, Garnotel E, Raoult D, Chaudet H. Insights Into Subspecies Discrimination Potentiality From Bacteria MALDI-TOF Mass Spectra by Using Data Mining and Diversity Studies. Front Microbiol. 2020; 13;11:1931.

[15]. Torres I, Giménez E, Vinuesa V, Pascual T, Moya JM, Alberola J et al. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) proteomic profiling of cerebrospinal fluid in the diagnosis of enteroviral meningitis: a proof-of-principle study. Eur J Clin Microbiol Infect Dis. 2018; 37:2331–39.

[16]. Chabriere E, Bassène H, Drancourt M, Sokhna C. MALDI-TOF MS and point of care are disruptive diagnostic tools in Africa. New Microbes New Infect. 2018; 26:S83–S88.

[17]. Laroche M, Bérenger J-M, Delaunay P, Charrel R, Pradines B, Berger F et al. Medical Entomology: A Reemerging Field of Research to Better Understand Vector-Borne Infectious Diseases. Clin Infect Dis. 2017; 65:S30–S38.

Table 1: Characteristics of the population in the study with the p-value resulting from the comparison of the three groups

Table 2: Sensitivity and specificity using the different number of differentially expressed peptide peaks between the tuberculosis cases and the non-tuberculosis controls

Table 3: Sensitivity and specificity using the different number of differentially expressed peptide peaks between the pulmonary tuberculosis cases and the lower respiratory tract infections

Figure 1: Box plots of the 25 top ranking peaks differentially expressed between the pulmonary tuberculosis and the non-TB control groups. Most of the differentially expressed protein peaks had higher intensity levels in the PTB group than in the non-TB control group

Figure 2: Box plots of the 10 top ranking peaks differentially expressed between the pulmonary tuberculosis and the non-TB LRTI groups. Most of the differentially expressed protein peaks had lower intensity levels in the PTB group than in the non-TB LRTI group

Figure 1: Box plots of the 25 top ranking peaks differentially expressed between the pulmonary tuberculosis and the non-TB control groups. Most of the differentially expressed protein peaks had higher intensity levels in the PTB group than in the non-TB control group



Figure 2: Box plots of the 10 top ranking peaks differentially expressed between the pulmonary tuberculosis and the non-TB LRTI groups. Most of the differentially expressed protein peaks had lower intensity levels in the PTB group than in the non-TB LRTI group



Characteristics	Pulmonary tuberculosis case N= 35 (%)	Non-tuberculosis controls N= 31 (%)	Non-tuberculosis lower respiratory tract infection N= 36 (%)	p-value	
Gender					
Male	30 (85.7)	22 (71.0)	27 (75.0)	0.00	
Female	5 (14.3)	9 (29.0)	9 (25.0)	0.33	
Age, mean (years) ± SD	37.5±15.0	40.1±18.8	59.3±16.8		
Age range	15 - 78	15 - 85	25 - 87	<0.000	
Chronic respiratory disease	4 (11.4)	7 (22.6)	15 (41.7)	0.01	
Diabetic	1 (2.9)	2 (6.5)	6 (16.7)	0.1	
Cardiovascular disease	4 (11.4)	6 (19.4)	17 (47.2)	0002	
Cancer	0	4 (12.9)	5 (13.9)	0.08	
Immunosuppressed	1 (2.9)	9 (29.0)	9 (25.0)	0.01	
Smoker	19 (54.3)	13 (41.9)	18 (50.0)	0.60	
Alcoholism	12 (34.3)	3 (9.7)	6 (16.7)	0.04	
Clinical symptoms					
General altered state	26 (74.3)	12 (38.7)	19 (52.8)	0.01	
Fever	18 (51.4)	13 (41.9)	29 (80.6)	0.003	
Night sweats	15 (42.9)	6 (19.4)	8 (22.2)	0.06	

Table 1: Characteristics of the population in the study with the p-value resulting from the comparison of the three groups

Weight loss	21 (60.0)	6 (19.4)	6 (18.2)	<0.000
Cough	31 (88.6)	15 (48.4)	34 (94.4)	<0.000
Haemoptysis	10 (28.6)	5 (16.1)	6 (16.7)	0.35
Chest pain	9 (25.7)	4 (12.9)	5 (13.9)	0.30
Dyspnoea	10 (28.6)	3 (9.7)	20 (55.6)	<0.000

Table 2: Sensitivity and specificity using the different number of differentially expressed peptide peaks between the tuberculosis cases and the non-tuberculosis controls

					Sensitivity	Specificity	Positive	Negative	
Diagnosis	True	False	False	True	(07.)	(07.)	predictive value	predictive value	Youden
model	positive	negative	positive	negative	(%)	(%)	(%)	(%)	index
					[95% CI]	[95% CI]	[95% CI]	[95% CI]	
					94.0	52.0	69.0	89.0	
All peaks	33	2	15	16	[81.0–99.0]	[33.0–70.0]	[54.0-81.0]	[48.0–89.0]	0.46
					83.0	48.0	64.0	71.0	
30 peaks	29	6	16	15	[66.0–93.0]	[30.0–67.0]	[49.0–78.0]	[48.0–89.0]	0.31
					97.0	77.0	83.0	96.0	
25 peaks	34	1	7	24	[85.0–100]	[59.0–90.0]	[68.0–93.0]	[80.0-100]	0.75
					91.0	77.0	82.0	89.0	
20 peaks	32	3	7	24	[77.0–98.0]	[59.0–90.0]	[66.0–92.0]	[71.0-98.0]	0.69

15 peaks	37	3	7	24	91.0	77.0	82.0	89.0	0.60	
15 peaks	52	5	7	24	[77.0–98.0]	[59.0–90.0]	[66.0–92.0]	[71.0-98.0]	0.09	
10 maaka	21	4	12	10	89.0	61.0	72.0	83.0	0.5	
10 peaks	31	31 4 12	12	12 19	[73.0–97.0]	[42.0–78.0]	[56.0-85.0]	[61.0–95.0]	0.3	

 Table 3: Sensitivity and specificity using the different number of differentially expressed peptide peaks between the pulmonary

 tuberculosis cases and the lower respiratory tract infections

					Sensitivity	Specificity	Positive	Negative	
Diagnosis	True	False	False	True	(%)	(%)	predictive value	predictive value	Youden
model	positive	negative	positive	negative	(%)	(%)	(%)	(%)	index
					[95% CI]	[95% CI]	[95% CI]	[95% CI]	
					54.0	78.0	58.0	54.0	
All peaks	19	16	8	28	[37.0–71.0]	[61.0–90.0]	[39.0–75.0]	[37.0–71.0]	0.32
30 peaks	22	13	1	35	63.0	97.0	87.0	78.0	0.60
I					[45.0–79.0]	[85.0–100]	[70.0–96.0]	[62.0–90.0]	
25 peaks	24	11	4	32	69.0	89.0	57.0	88.0	0.57
25 peaks	27	11	Ŧ	52	[51.0-83.0]	[74.0–97.0]	[43.0–69.0]	[47.0–100]	0.57
20 peaks	22	13	3	33	63.0	92.0	79.0	96.0	0.54

				[45.0–79.0]	[78.0–98.0]	[64.0–90.0]	[80.0-100]	
10	16	0	26	54.0	100	85.0	82.0	0.54
19	10	0	30	[37.0–71.0]	[90.0–100]	[69.0–95.0]	[65.0–93.0]	0.34
28	7	4	20	80.0	89.0	86.0	72.0	0.60
28	7	4	52	[63.0–92.0]	[74.0–97.0]	[67.0–96.0]	[56.0-85.0]	0.09
	19 28	19 16 28 7	19 16 0 28 7 4	19 16 0 36 28 7 4 32	$\begin{bmatrix} 45.0-79.0 \end{bmatrix}$ $\begin{bmatrix} 19 & 16 & 0 & 36 \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ &$	$\begin{bmatrix} 45.0-79.0 \end{bmatrix} \begin{bmatrix} 78.0-98.0 \end{bmatrix}$ $\begin{bmatrix} 19 & 16 & 0 & 36 \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ &$	$\begin{bmatrix} 45.0-79.0 \end{bmatrix} \begin{bmatrix} 78.0-98.0 \end{bmatrix} \begin{bmatrix} 64.0-90.0 \end{bmatrix}$ $\begin{bmatrix} 19 \\ 16 \\ 0 \\ 36 \end{bmatrix} \begin{bmatrix} 54.0 \\ 100 \end{bmatrix} \begin{bmatrix} 100 \\ 85.0 \end{bmatrix}$ $\begin{bmatrix} 37.0-71.0 \end{bmatrix} \begin{bmatrix} 90.0-100 \end{bmatrix} \begin{bmatrix} 69.0-95.0 \end{bmatrix}$ $\begin{bmatrix} 80.0 \\ 89.0 \end{bmatrix} \begin{bmatrix} 89.0 \\ 89.0 \end{bmatrix} \begin{bmatrix} 86.0 \\ 86.0 \end{bmatrix}$ $\begin{bmatrix} 28 \\ 7 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} \begin{bmatrix} 22 \\ 23 \end{bmatrix} \begin{bmatrix} 63.0-92.0 \end{bmatrix} \begin{bmatrix} 74.0-97.0 \end{bmatrix} \begin{bmatrix} 67.0-96.0 \end{bmatrix}$	$\begin{bmatrix} 45.0-79.0 \end{bmatrix} \begin{bmatrix} 78.0-98.0 \end{bmatrix} \begin{bmatrix} 64.0-90.0 \end{bmatrix} \begin{bmatrix} 80.0-100 \end{bmatrix}$ $\begin{bmatrix} 19 \\ 16 \end{bmatrix} \begin{bmatrix} 0 \\ 36 \end{bmatrix} \begin{bmatrix} 54.0 \\ 100 \end{bmatrix} \begin{bmatrix} 100 \\ 85.0 \end{bmatrix} \begin{bmatrix} 85.0 \\ 82.0 \end{bmatrix}$ $\begin{bmatrix} 37.0-71.0 \end{bmatrix} \begin{bmatrix} 90.0-100 \end{bmatrix} \begin{bmatrix} 69.0-95.0 \end{bmatrix} \begin{bmatrix} 65.0-93.0 \end{bmatrix}$ $\begin{bmatrix} 28 \\ 7 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} \begin{bmatrix} 2 \\ 32 \end{bmatrix} \begin{bmatrix} 80.0 \\ 89.0 \end{bmatrix} \begin{bmatrix} 89.0 \end{bmatrix} \begin{bmatrix} 80.0 \end{bmatrix} \begin{bmatrix} 69.0-95.0 \end{bmatrix} \begin{bmatrix} 65.0-93.0 \end{bmatrix}$ $\begin{bmatrix} 63.0-92.0 \end{bmatrix} \begin{bmatrix} 74.0-97.0 \end{bmatrix} \begin{bmatrix} 67.0-96.0 \end{bmatrix} \begin{bmatrix} 56.0-85.0 \end{bmatrix}$