

UC Berkeley

UC Berkeley Previously Published Works

Title

A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases.

Permalink

<https://escholarship.org/uc/item/5p58k729>

Journal

Tuberculosis, 95(4)

Authors

Laux da Costa, Lucas

Delcroix, Melaine

Dalla Costa, Elis

et al.

Publication Date

2015-07-01

DOI

10.1016/j.tube.2015.04.008

Peer reviewed



HHS Public Access

Author manuscript

Tuberculosis (Edinb). Author manuscript; available in PMC 2016 July 01.

Published in final edited form as:

Tuberculosis (Edinb). 2015 July ; 95(4): 421–425. doi:10.1016/j.tube.2015.04.008.

A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases

Lucas Laux da Costa^{a,b,**}, Melaine Delcroix^{c,**}, Elis R. Dalla Costa^a, Isaías V. Prestes^d, Mariana Milano^{a,b}, Steve S. Francis^e, Gisela Unis^f, Denise R. Silva^{b,g}, Lee W. Riley^c, and Maria L. R. Rossetti^{a,h}

Lucas Laux da Costa: lauxlc@yahoo.com.br; Melaine Delcroix: mdelcroix@berkeley.edu; Elis R. Dalla Costa: dallacostaer@gmail.com; Isaías V. Prestes: ppgepid@ufrgs.br; Mariana Milano: marianamilano@hotmail.com; Steve S. Francis: ssfrancis@berkeley.edu; Gisela Unis: giunis@terra.com.br; Denise R. Silva: denise.rossato@terra.com.br; Lee W. Riley: lwiley@berkeley.edu; Maria L. R. Rossetti: mrossett@terra.com.br

^aFundação Estadual de Produção e Pesquisa em Saúde, Avenida Ipiranga, 5400, Porto Alegre, RS 90610-000, Brazil

^bUniversidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500, Porto Alegre, RS 91501-970, Brazil

^cDivision of Infectious Disease and Vaccinology, School of Public Health, University of California, 530 Li Ka Shing, Berkeley, CA 94720, USA

^dPrograma de Pós Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, Rua Ramiro Barcelos, 2400, Porto Alegre, RS 90035-003, Brazil

^eDivision of Epidemiology, School of Public Health, University of California, 50 University Hall, Berkeley, CA 94720, USA

^fHospital Sanatório Partenon, Avenida Bento Gonçalves, 3722, Porto Alegre, RS 90650-001, Brazil

^gHospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos, 2350, Porto Alegre, RS 90035-903, Brazil

^hUniversidade Luterana do Brazil, Avenida Farroupilha, 8001, Bairro São José, Canoas, RS 92425-900, Brazil

Abstract

© 2015 Published by Elsevier Ltd.

Corresponding author: mdelcroix@berkeley.edu; 530 Li Ka Shing, Berkeley, CA 94720; Tel: +1 510-643-2949; fax: +1 510-643-4966.

**These authors contributed equally to this work.

Conflict of interest statement: None of the Authors that contributed to this work have any conflict of interest to declare.

Ethical approval: The study was accepted by the ethical committees of the Prefeitura de Porto Alegre (IRB approval 630), the Hospital de Clínicas in Porto Alegre (IRB approval 110201) and the Fundação Estadual de Produção e Pesquisa em Saúde (IRB approval 03/2011).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The goal of this study was to identify a host gene signature that can distinguish tuberculosis (TB) from other pulmonary diseases (OPD). We conducted real-time PCR on whole blood samples from patients in Brazil. TB and OPD patients (asthma and non-TB pneumonia) differentially expressed granzyme A (*GZMA*), guanylate binding protein 5 (*GBP5*) and Fc gamma receptor 1A (*CD64*). Receiver operating characteristic, tree classification and random forest analyses were applied to evaluate the discriminatory power of the three genes and find the gene panel most predictive of patients' disease classification. Tree classification produced a model based on *GBP5* and *CD64* expression. In random forest analysis, the combination of the three genes provided a robust biosignature to distinguish TB from OPD with 95% specificity and 93% sensitivity. Our results suggest that *GBP5* and *CD64* in tandem may be the most predictive combination. However, *GZMA* contribution to the prediction model requires further investigation. Regardless, these three genes show promise as a rapid diagnostic marker separating TB from OPD.

Keywords

tuberculosis; diagnostics; *GBP5*; *CD64*; *GZMA*; biomarkers

1. Introduction

Despite increasing treatment success rates, tuberculosis (TB) continues to spread worldwide. The World Health Organization (WHO) estimates that 8.6 million people developed TB and 1.3 million died from the disease in 2012 (WHO, 2013). Lack of accurate and rapid test to diagnose TB remains an important obstacle to TB control. The most common clinical dilemma encountered in TB diagnosis is the differentiation of pulmonary TB from other common lung diseases. When a patient is found to have an abnormal chest X-ray, TB is not necessarily on the top of the list of differential diagnosis in most settings.

Unbiased microarray gene expression profiling of whole blood cells has provided candidate biosignatures to discriminate TB patients from healthy donors (Berry et al., 2010; Jacobsen et al., 2007; Lu et al., 2011; Maertzdorf et al., 2011; Ottenhoff et al., 2012). However, these studies that compare TB to healthy controls mostly reveal a general state of persistent inflammation and not TB per se (Weiner, Maertzdorf, & Kaufmann, 2013).

Recently, several groups have sought TB-specific biomarkers by comparing them in patients with TB versus other inflammatory pulmonary diseases (Bloom et al., 2013; Kaforou et al., 2013; Koth et al., 2011; Maertzdorf et al., 2012). Despite overlapping gene up-regulation in proinflammatory pathways and general immunopathological mechanisms, they could identify transcript signatures to distinguish TB from other diseases, including in HIV infected and uninfected subjects (Kaforou et al., 2013). Here, we selected a restricted set of genes to assess their ability to differentiate TB from other pulmonary diseases (OPD) patients. These genes were previously identified by Maertzdorf et al. (Maertzdorf et al., 2011) as a biosignature to discriminate TB and healthy latently infected (LTBI) subjects: granzyme A (*GZMA*), guanylate binding protein 5 (*GBP5*), Fc gamma receptor 1A (*CD64*), Fc gamma receptor 1B (*FCGR1B*) and lactotransferrin (*LTF*). Maertzdorf et al. (2011) did not include OPD patients in their study, but since then, *CD64*, *FCGR1B* and *LTF* were

shown to be differentially expressed in TB versus other lung diseases (Bloom et al., 2013; Joosten, Fletcher, & Ottenhoff, 2013; Maertzdorf et al., 2012).

2. Materials and Methods

2.1. Study Participants

Subjects were recruited between March 1, 2011 and March 30, 2013. Written informed consent was obtained from all participants. Our cohort included 27 patients with active tuberculosis (TB), 27 healthy donors with latent *M. tuberculosis* infection (LTBI), 25 healthy non-infected donors (NIDs), and 22 patients suffering from other pulmonary diseases (OPD)--14 patients with asthma and eight patients with streptococcal pneumonia (PN). All subjects were older than 18, and responded to a standardized questionnaire (Table 1).

TB patients were recruited at the Sanatório Partenon Hospital, Porto Alegre, Brazil. All samples were collected prior to TB treatment from patients without prior TB infection. Patients were confirmed to have pulmonary TB by chest x-ray, sputum smear microscopy and culture. Healthy LTBI donors were recruited at the Unidade Básica de Saúde Restinga in Porto Alegre. LTBI subjects were household contacts of TB patients whose tuberculin skin test (TST) showed an induration >5 mm, a normal chest X-ray, and no clinical symptoms of TB or OPD. NIDs and patients with OPD were recruited at the Clinical Hospital in Porto Alegre. NIDs had a negative TST (< 5 mm), no clinical symptoms of TB or OPD, and a normal chest x-ray. OPD patients had a negative TST, normal chest X-ray for asthma and abnormal for PN. Gram staining and culture of sputum samples confirmed the diagnosis of PN due to *Streptococcus pneumoniae*. Immunological tests for HIV/HCV/HBV were negative in all participants.

2.2. RNA Extraction and Real-time PCR Quantification

We collected 2.5 mL peripheral whole blood in a PAXgene blood RNA tube (PreAnalytiX) for every donor. Total RNA was extracted with the PAXgene Blood RNA kit (Qiagen). cDNA was generated by reverse transcription with Superscript III according to manufacturer's instructions (Life Technologies) and quantified by real-time PCR with SsoAdvanced SYBR Green Supermix (BioRad) on the CFX96 Touch Real-Time PCR Detection System (BioRad). Primer sequences were those published by Maertzdorf et al. (Maertzdorf et al., 2011).

2.3. Statistical Analysis

Scatter plots were generated with the ggplot2 package in R (Team, 2013). Receiver operating characteristic (ROC), classification tree and random forest (randomForest) analyses were performed with the pROC, OptimalCutpoints (Youden method), tree and randomForest (Liaw & Wiener, 2002) packages, respectively, in R. No separate test set was needed for random forest analysis as this algorithm leaves one third of the cases out (out-of-bag) when constructing each tree.

3. Results

We assessed the gene expression level of *CD64*, *FCGR1B*, *GZMA*, *GBP5* and *LTF* by real-time PCR in whole blood samples from 25 NIDs, 27 LTBI, 27 TB, 14 asthma and eight PN donors (Table 1). *B2M* was chosen as reference gene based on Maertzdorf et al.'s study (Maertzdorf et al., 2011). There was a difference in age distribution between the study groups, but we found no evidence to suggest gene expression levels correlated with the donor's age. Although we did find evidence that suggests *FCGR1B* and *LTF* expression is affected by gender, a small sample size precluded detailed stratification (data not shown). The difference in *FCGR1B* expression in female TB and OPD patients had a *p*-value of less than 0.001 while no difference in expression was seen in male subjects. Female TB and OPD patients did not show any difference in *LTF* expression, whereas male TB and OPD patients had a *p*-value less than 0.05 for *LTF*. Thus, we excluded *FCGR1B* and *LTF* in our predictive model and continued our analysis with *CD64*, *GZMA* and *GBP5*. These three genes showed significant difference in gene expression between TB and OPD patients with a *p*-value less than 0.01 for *CD64* and less than 0.001 for *GZMA* and *GBP5* (Figure 1).

ROC methodology was applied to evaluate the individual discriminatory ability of the three genes. The values of the area under the curve (AUC) and the optimal cut-off points are shown in Figure 2a. *GBP5* expression was the best parameter to discriminate between TB and OPD with an AUC of 0.924 with a 95% confidence interval (CI) of 0.839–1.000. *GZMA* AUC was equal to 0.889 (0.769–1.000) while *CD64* performed with the lowest AUC at 0.778 (0.640–0.916).

Thereafter, we subjected the three genes to decision tree analysis to find the ideal gene combination to optimize the discrimination between TB and OPD (Figure 2b). This analysis showed that a combination of *GBP5* and *CD64* provided the best predictive capacity with 47 of 49 individuals being correctly classified. The sensitivity of this combination was 93%, as 25 of the 27 TB patients were correctly identified. The specificity reached 100%, as all of 22 OPD individuals were correctly classified as non-TB. The first decision node was based on *GBP5* expression and achieved 93% sensitivity. *CD64* expression level improved specificity from 86% to 100% by identifying misclassified non-TB cases. We then ran random forest analysis to confirm that the two-gene combination obtained by decision tree would also be identified by a different classification methodology (Figure 2c). The combination of the three genes gave the highest accuracy in distinguishing TB from OPD patients. It gave a specificity of 95% and a sensitivity of 93% to distinguish TB from non-TB cases. *GBP5* showed the highest discriminating power with an importance value of 24, followed by *GZMA* (20) and *CD64* (9). When running ROC analysis on the random forest votes, we obtained an AUC of 0.955 (0.885–1.000) for the three genes. The combination of *GBP5* and *CD64* did not perform as well as the three-gene combination to discriminate TB versus non-TB patients (82% specificity and 85% sensitivity). Interestingly, *GZMA* alone showed a sensitivity of 96% with only one of 27 TB patients misclassified (although the specificity was 77%).

To further validate *GBP5*, *GZMA* and *CD64* discriminative power, we applied random forest analysis on the previously published data from Bloom et al. microarray study independently

of our trained dataset (Bloom et al., 2013). We extracted the values on *GBP5*, *CD64* and *GZMA* expression in TB and PN patients and ran them through the random forest algorithm for training and out-of-bag testing. We found that the panel of *CD64*, *GZMA* and *GBP5* had a specificity of 84% and a sensitivity of 94% to distinguish TB from non-TB cases (Figure 2c). Two of 35 TB patients were misclassified and three of 19 PN patients were identified as TB patients. *GBP5* importance value reached 25 followed by *CD64* (8) and *GZMA* (4). ROC analyses on the random forest votes produced an AUC of 0.955 (0.899–1.000) for the three-gene combination. *GBP5* alone or in tandem with *CD64* or *GZMA* produced similar specificity, sensitivity and AUC values as the three-gene panel, confirming *GBP5* discriminative power on the Bloom et al. dataset.

For all three genes, gene expression levels were found significantly different in TB versus NIDs and TB versus LTBI with a *p*-value less than 0.001 (Figure 1). Although at a lesser degree, OPD subjects also showed a differential expression of *CD64* compared to LTBI and NIDs. *GBP5* and *GZMA* were expressed at similar levels in NIDs, LTBI and OPD donors.

4. Discussion

Clinically relevant TB biomarkers need to be able to distinguish pulmonary TB and non-TB pulmonary diseases. Otherwise, such tests would be no more specific than a chest x-ray. Here, we wished to be able to distinguish TB from other common pulmonary diseases. Asthma represents a chronic non-infectious inflammatory airways disease. Non-TB pneumonias represent the most common infectious pulmonary disease that healthcare providers need to rapidly distinguish from TB, when patients seek health care with symptoms of lung disease and an abnormal chest x-ray.

Blood transcriptomic studies have identified gene signatures that can discriminate LTBI individuals from TB patients (Berry et al., 2010; Jacobsen et al., 2007; Maertzdorf et al., 2011). However, the identified gene expression profiles mainly reflect a state of persistent activation of the immune system (Maertzdorf et al., 2012; Weiner et al., 2013). When interrogating the transcriptome from patients with different pulmonary conditions, studies revealed an overlap with the previously identified gene expression patterns attributed to TB (Bloom et al., 2013; Kaforou et al., 2013; Koth et al., 2011; Maertzdorf et al., 2012). In this study, the rationale to work with a restricted number of genes was that a larger set of biomarkers would likely encompass gene expression profiles shared by several diseases. We used RT-PCR to assess the expression level of five genes, *GZMA*, *GBP5*, *CD64*, *FCGR1B* and *LTF*. These genes were previously identified by Maertzdorf et al. as they formed a biosignature capable of distinguishing between TB and LTBI with a sensitivity and a specificity of 94% (30/32 donors) and 97% (33/34 donors), respectively (Maertzdorf et al., 2011). Our results confirmed that *GZMA*, *GBP5* and *CD64* were differentially expressed in TB versus NIDs and TB versus LTBI donors. More interestingly, we showed the discriminative power of the three genes in distinguishing TB and OPD patients.

FCGR1 was reported as one of the most differentially expressed genes in TB patients and LTBI donors (Maertzdorf et al., 2011). We could not distinguish *FCGR1B* expression between TB and OPD patients as gender seemed to alter its expression in our cohort.

Likewise, *LTF* expression profile varied in women versus men, precluding us from using this gene as a marker. We found that the gene expression levels of *GBP5*, *GZMA* and *CD64* were significantly different in TB versus OPD patients. *GZMA* had a higher Ct value in TB patients compared with the other patients' groups, indicating a lower expression of *GZMA* in TB patients, as reported by Maertzdorf et al. (Maertzdorf et al., 2011). On the other hand, *GBP5* and *CD64* Ct values were at their lowest in TB patients, reflecting a higher expression in this patients' group.

We performed three types of analysis to determine the discriminatory power of *GBP5*, *GZMA* and *CD64* genes: ROC, tree and random forest methodologies. In all analyses, *GBP5* was consistently identified as the strongest predictor of TB versus non-TB cases. In the model provided by the tree software, *CD64* had a minor contribution but did optimize the classification of the patients after a first split based on *GBP5* expression. All OPD patients were correctly identified and only two TB patients were misclassified.

Using random forest analysis, we were able to classify the study participants as belonging to either TB or OPD groups with a high accuracy and the combination of all three genes, *GBP5*, *GZMA* and *CD64*, formed the strongest biosignature with 95% specificity (21/22 OPD) and 93% sensitivity (25/27 TB). While the tree model did not include *GZMA*, this gene was identified as a strong predictor by random forest analysis. The importance of *GBP5* as a biomarker was confirmed by random forest analysis run on the Bloom et al. study dataset: *GBP5* alone produced similar sensitivity, specificity values as the three-gene panel, or *GBP5* in tandem with either *CD64* or *GZMA* to classify patients.

GBP5 and *GZMA* were shown to be differentially expressed by TB patients compared to asymptomatic controls (Maertzdorf et al., 2011) but this had not been validated against patients with other diseases (Bloom et al., 2013; Kaforou et al., 2013; Maertzdorf et al., 2012). *GBP5* was one of the genes up-regulated in sarcoidosis lung tissue when compared to healthy donors in Koth et al.'s study (Koth et al., 2011). However, Koth et al. did not identify *GBP5* to vary between sarcoidosis and TB when they used previously published blood gene expression datasets from TB cohorts (Berry et al., 2010). As reported by Blankley et al. (Blankley et al., 2014), studies designed to directly compare diseases reveal different signatures than studies based on previously published dataset. Also, when comparing biomarker studies, many factors such as population ethnicity and co-infections, may affect results (Joosten et al., 2013).

CD64 was included in the panel of 144 transcripts identified by Bloom et al. as differentially expressed in TB versus other lung diseases (Bloom et al., 2013). In our study, TB patients showed a higher *CD64* expression than OPD patients. In their meta-analysis of TB microarray datasets, Joosten et al. found an increased expression of genes associated with Fc receptor signaling, including *CD64* in TB patients (Joosten et al., 2013). *LTF* and *FCGR1B* were also identified as differentially expressed genes in TB versus OPD patients in Maertzdorf et al. (Maertzdorf et al., 2012) and Bloom et al. (Bloom et al., 2013) reports, respectively, but could not be confirmed as markers in our study.

One limitation of this study is small sample size, which explains the discrepancy in *GZMA* contribution to patients' classification. It is therefore difficult to determine the exact combination of the *GBP5*, *GZMA* and *CD64* genes to maximize predictive power. Our results suggest that *GBP5* and *CD64* in tandem may be the most predictive combination. Without further studies we cannot rule out the contribution of *GZMA* to differentiate TB from OPD. Regardless, these three genes show promise as a diagnostic marker to distinguish TB from OPD and require further investigation.

Acknowledgments

Funding statement: This work was supported by the Fogarty International Center at the National Institutes of Health [grant number U2RTW006885] and the UBS Optimus Foundation.

Abbreviations

<i>CD64</i>	Fc gamma receptor 1A
<i>FCGR1B</i>	Fc gamma receptor 1B
<i>GBP5</i>	guanylate binding protein 5
<i>GZMA</i>	granzyme A
LTBI	healthy latently infected
LTF	lactotransferrin
NID	healthy non-infected donor
OPD	other pulmonary diseases
PN	pneumonia
TB	tuberculosis

References

- Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, O'Garra A. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*. 2010; 466(7309): 973–977.10.1038/nature09247 [PubMed: 20725040]
- Blankley S, Berry MP, Graham CM, Bloom CI, Lipman M, O'Garra A. The application of transcriptional blood signatures to enhance our understanding of the host response to infection: the example of tuberculosis. *Philos Trans R Soc Lond B Biol Sci*. 2014; 369(1645):20130427.10.1098/rstb.2013.0427 [PubMed: 24821914]
- Bloom CI, Graham CM, Berry MP, Rozakeas F, Redford PS, Wang Y, O'Garra A. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PLoS One*. 2013; 8(8):e70630.10.1371/journal.pone.0070630 [PubMed: 23940611]
- Jacobsen M, Repsilber D, Gutschmidt A, Neher A, Feldmann K, Mollenkopf HJ, Kaufmann SH. Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J Mol Med (Berl)*. 2007; 85(6):613–621.10.1007/s00109-007-0157-6 [PubMed: 17318616]
- Joosten SA, Fletcher HA, Ottenhoff TH. A helicopter perspective on TB biomarkers: pathway and process based analysis of gene expression data provides new insight into TB pathogenesis. *PLoS One*. 2013; 8(9):e73230.10.1371/journal.pone.0073230 [PubMed: 24066041]

- Kaforou M, Wright VJ, Oni T, French N, Anderson ST, Bangani N, Levin M. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med.* 2013; 10(10):e1001538.10.1371/journal.pmed.1001538 [PubMed: 24167453]
- Koth LL, Solberg OD, Peng JC, Bhakta NR, Nguyen CP, Woodruff PG. Sarcoidosis blood transcriptome reflects lung inflammation and overlaps with tuberculosis. *Am J Respir Crit Care Med.* 2011; 184(10):1153–1163.10.1164/rccm.201106-1143OC [PubMed: 21852540]
- Liaw A, Wiener M. Classification and Regression by randomForest. *R News.* 2002; 2(3):18–22.
- Lu C, Wu J, Wang H, Wang S, Diao N, Wang F, Zhang W. Novel biomarkers distinguishing active tuberculosis from latent infection identified by gene expression profile of peripheral blood mononuclear cells. *PLoS One.* 2011; 6(8):e24290.10.1371/journal.pone.0024290 [PubMed: 21904626]
- Maertzdorf J, Repsilber D, Parida SK, Stanley K, Roberts T, Black G, Kaufmann SH. Human gene expression profiles of susceptibility and resistance in tuberculosis. *Genes Immun.* 2011; 12(1):15–22.10.1038/gene.2010.51 [PubMed: 20861863]
- Maertzdorf J, Weiner J 3rd, Mollenkopf HJ, Network TB, Bauer T, Prasse A, Kaufmann SH. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc Natl Acad Sci U S A.* 2012; 109(20):7853–7858.10.1073/pnas.1121072109 [PubMed: 22547807]
- Ottenhoff TH, Dass RH, Yang N, Zhang MM, Wong HE, Sahiratmadja E, Hibberd ML. Genome-wide expression profiling identifies type 1 interferon response pathways in active tuberculosis. *PLoS One.* 2012; 7(9):e45839.10.1371/journal.pone.0045839 [PubMed: 23029268]
- Team, R. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2013.
- Weiner J, Maertzdorf J, Kaufmann SH. The dual role of biomarkers for understanding basic principles and devising novel intervention strategies in tuberculosis. *Ann N Y Acad Sci.* 2013; 1283:22–29.10.1111/j.1749-6632.2012.06802.x [PubMed: 23181737]
- WHO. Global tuberculosis report 2013. World Health Organization; Geneva: 2013.

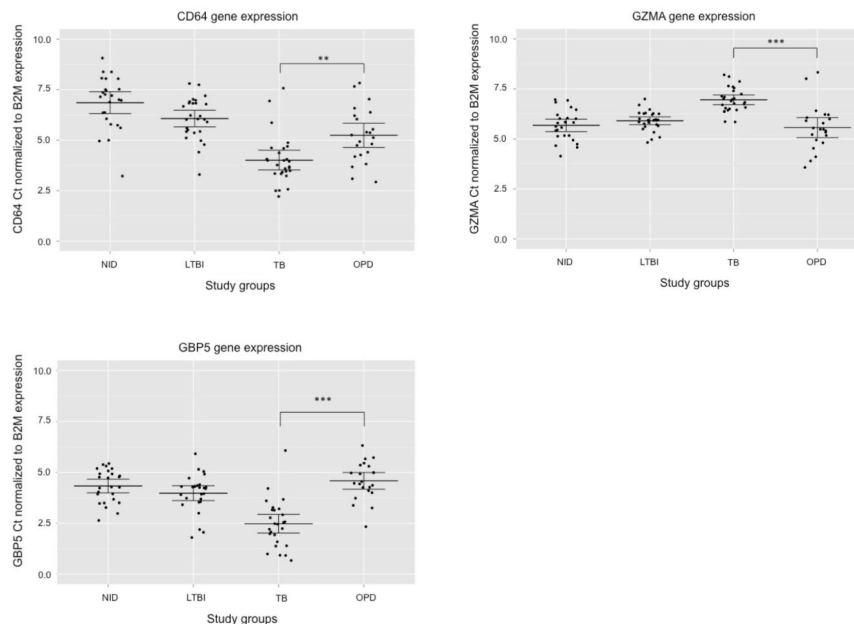
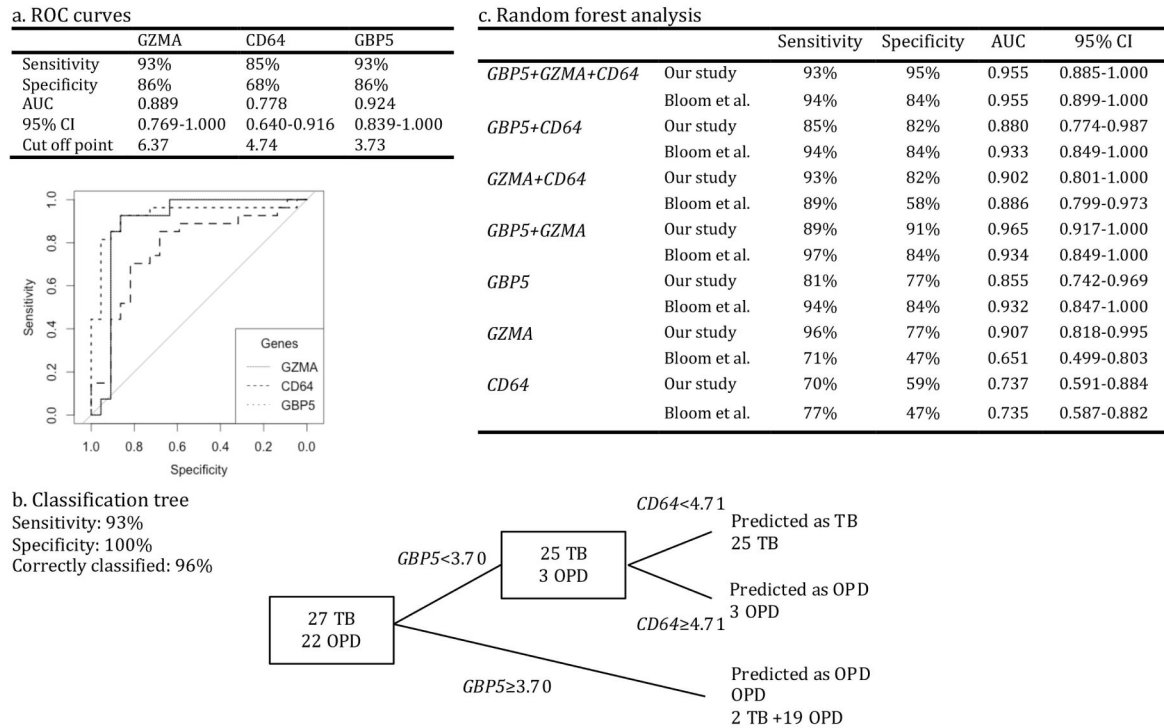


Figure 1. Differences in gene expression levels between NIDs, LTBI, TB and OPD donors. Scatter dot plots for *CD64*, *GZMA* and *GBP5* represent the gene expression for each individual donor. Ct values from real-time PCR were normalized to the internal reference gene *B2M*. The mean and 95% confidence interval are indicated. ** indicates a p -value < 0.01 ; *** indicates a p -value < 0.001 . NID, healthy non-infected donors; LTBI, healthy donors latently infected with *M. tuberculosis*; TB, tuberculosis patients; OPD, patients with other pulmonary disease than TB (asthma and pneumonia).

**Figure 2.**

Analyses performed to evaluate the discriminatory ability of *GBP5*, *GZMA* and *CD64* genes and find the optimal gene combination to discriminate between TB and OPD groups. a, ROC analysis identified *GBP5* as the strongest discriminatory gene. b, Classification tree is shown based on the gene expression of *GBP5* and *CD64*. The “tree” software did not include *GZMA*. The sensitivity and specificity of the two-gene panel was 93% and 100%, respectively, and 96% individuals were correctly classified as TB or OPD. c, Sensitivity and specificity values obtained from random forest analysis of *GBP5*, *GZMA* and *CD64* combinations applied to our study and Bloom’s (Bloom et al., 2013) data. AUC values (with 95% confidence interval) obtained by running ROC analysis on random forest votes are also given. AUC, area under curve; CI, confidence interval; OPD, patients with other inflammatory pulmonary diseases (asthma and pneumonia); ROC, receiver operating characteristic; TB, tuberculosis patients.

Table 1

Demographic and diagnostic characteristics of study groups.

	NID	L/TBI	TB	Asthma	PN
Number of subjects	25	27	27	22	22
Females/Males	12/13	15/12	9/18	15/7	15/7
Age	32.6 (12.5)	42.1 (15.8)	34.4 (13.6)	55.0 (10.2)	55.0 (10.2)
Range	20–62	18–65	19–65	26–65	26–65
Sputum smear	NA	NA	Positive	NA	NA
Sputum culture	NA	NA	Positive	NA	NA
Infectious organism	NA	NA	<i>Mycobacterium tuberculosis</i>	NA	<i>Streptococcus pneumoniae</i>
TST	< 5 mm	> 5 mm	NA	< 5 mm	< 5 mm
Chest X-Ray					
Normal	25 (100%)	27 (100%)	1 (4%)	14 (100%)	0
Cavitation	0	0	1 (4%)	0	0
Consolidation	0	0	1 (4%)	0	8 (100%)
Cavitation and consolidation	0	0	24 (88%)	0	0

NID, healthy non-infected donors; L/TBI, healthy donors latently infected with *M. tuberculosis*; TB, tuberculosis patients; PN, patients with pneumonia; NA, non applicable; TST, tuberculin skin test.