

TITLE: Diagnostic utility of in vivo expressed mycobacterial RNA transcripts in pleural fluid for the differential diagnosis of tuberculous pleuritis.

Prabhdeep Kaur¹, Sumedha Sharma¹, Sudhanshu Abhishek Sinha¹, Ashutosh N. Aggarwal², Khushpreet Kaur¹, Rakesh Yadav³, Sunil Sethi³, Amanjit Bal⁴, Indu Verma^{1*}

1.) Department of Biochemistry, Postgraduate Institute of Medical Education and Research, Chandigarh, India.

2.) Department of Pulmonary Medicine, Postgraduate Institute of Medical Education and Research, Chandigarh, India.

3.) Department of Medical Microbiology, Postgraduate Institute of Medical Education and Research, Chandigarh, India.

7.) Department of Histopathology, Postgraduate Institute of Medical Education and Research, Chandigarh, India.

*CORRESPONDENCE: Dr. Indu Verma, Professor and Head, Department of Biochemistry, Postgraduate Institute of Medical Education and Research, Chandigarh, India. Email: induvermabio@gmail.com

ABSTRACT

Background: Tuberculous pleuritis (TBP), the extra-pulmonary manifestation of tuberculosis, is the second most common after TB lymphadenitis. Histopathology using biopsy samples is the most sensitive diagnostic procedure for TBP, however the biopsy acquisition is invasive. Therefore, better screening markers for diagnosis using pleural fluid are required. The pathogen biomarkers expressed at the site of infection may play a potential role in designing a newer diagnostic assay. Thus, the current study was planned to look for mycobacterial RNA biomarkers in TBP and to assess their diagnostic utility in pleural fluid. **Methods:** TBP suspects (n=261) were recruited in the current study. Out of these 45 suspects were excluded and the remaining (n=216) were divided into TBP (n=54) and non-TBP (n=162) groups based on composite reference standard. A whole genome microarray was carried using *M.tb* RNA from pleural biopsies of TB patients. The data was validated using qRT-PCR and the diagnostic utility of top two highly expressed genes was assessed in pleural fluid of using a real time RT-PCR assay. **Results:** Overall, 1856 genes were differentially expressed in microarray of which 1365 were upregulated and 491 were downregulated. After validation of microarray gene expression, two genes namely *Rv1586* and *Rv2543* were selected for assessment of their diagnostic utility in TBP. The combined analysis for the presence of either of genes in the pleural fluid led to identification of pleural TB patients with 79.6% sensitivity and 93.28% specificity. **Conclusion:** The transcripts of genes *Rv1586* and *Rv2543* holds potential for the development of a RNA based molecular diagnostic assay in pleural fluid of TBP patients.

INTRODUCTION:

Tuberculosis typically affects lungs, but can also affect other sites. [1]. Extra-pulmonary TB (EPTB) accounts for approximately 15-20% cases in HIV negative people, and 50% in HIV positive individuals [2]. Tuberculous pleuritis (TB pleuritis/TBP) is the most prevalent form of extra-pulmonary tuberculosis globally after lymph node TB [2,3]. The diagnosis of TB pleuritis is carried out using pleural fluid and biopsy acquired through thoracentesis and thoracoscopy respectively. Considering the paucibacillary nature of TBP, the conventional diagnostic tests including pleural fluid microscopy and culture offer low sensitivity [4-7]. Pleural fluid cytological findings are mostly diagnostic for malignant effusions [2,8,9]. Histopathological examination of pleural biopsy exhibits highest diagnostic potential; however, biopsy acquisition is an invasive procedure [10,11] and associated with post procedure complications [12-14]. The adenosine deaminase (ADA), interferon gamma (IFN- γ) and T cell based IFN- γ release assays (IGRAs) [15-19] are rapid and cost effective, but are nonspecific and the utilization of in-house detection procedures [2,15,18] and lack of a widely accepted cutoff are major constraints limiting their utilization [20]. The utility of Xpert MTB/RIF, the only molecular assay available for TBP is limited [2,21]. The improved version Ultra has better sensitivity and has been endorsed by WHO for TBP diagnosis [22-24]. However, these NAAT's are based on DNA and cannot differentiate live/dead bacilli which could be reason for false positive Xpert MTB/RIF results in previously treated TB patients [25]. To address these limitations, in the current study, an RNA-based diagnostic test is proposed for TBP. For an RNA based assay the transcripts expressed *in vivo* at the site of infection by the pathogen could be important targets. Mycobacterial gene expression is known to be altered under different growth conditions as well as *in vivo* at different sites [26-32]. The study is based on the hypothesis that *in vivo* highly expressed mycobacterial genes particularly during active disease could be important signatures for the development of a sensitive and specific molecular diagnostic assay. Thus, the present study was carried out to analyse the mycobacterial gene expression profile in biopsy samples of TB pleuritis and identify mycobacterial transcripts that can be utilized as potential candidates for RNA based molecular assay for the diagnosis of pleural TB using pleural fluid samples. Though, RNA based diagnostic assays are difficult to perform due to technical complexities but a widespread use of these assays in the COVID era has provided new hopes for use of such assays for TB diagnosis.

MATERIALS AND METHODS

Study subjects

The study subjects for the present study were recruited from Department of Pulmonary Medicine, PGIMER following the approval of study by Institutional Ethics Committee (IEC) (Ref no. INT/IEC/2016/2264). Patients with exudative pleural effusions, history of fever and/or respiratory symptoms for >2 weeks and referred for thoracentesis for diagnostic/therapeutic purpose were recruited as tuberculosis pleuritis (TBP) suspects (n=261) with prior written

informed consent. Patients already on anti-tubercular therapy (ATT) at time of recruitment or having positive HIV/HCV status or with any other co-morbidity were excluded from the study.

Sample collection: Pleural biopsy and pleural fluid samples were collected from study subjects using thoracoscopy and thoracentesis, respectively by trained physicians. All the samples were also subjected to routine diagnostic tests and the reports were prospectively collected. Clinical and demographic details were collected at the time of enrollment.

Diagnostic algorithm for categorization of TBP suspects: A composite reference standard (CRS) comprising of clinical findings (history of fever or respiratory symptoms along with radiological evidence), microscopy (presence of acid-fast bacilli in pleural fluid/ pleural biopsy), culture (in pleural fluid/ pleural biopsy), histopathology (pleural biopsy showing the presence of typical caseating granulomatous inflammation), Xpert MTB/RIF or Xpert MTB/RIF Ultra positivity, elevated ADA (>40 IU/L) and response to ATT (improvement in effusion after starting ATT) was utilized for diagnosis of TBP patients. The patients showing positivity with any one of the tests from CRS were categorized as TBP. The patients presenting with pleural effusions showing the presence of cytological or histological evidence of malignant cells/etiology of other respiratory infections in pleural fluid/pleural biopsy but not showing the presence of *M.tb* in pleural fluid/biopsy were classified as non TB pleuritis (Non TBP) patients.

Mycobacterial RNA isolation from pleural biopsy and pleural fluid

Pleural biopsy samples in RNA later were centrifuged at 3000g for 15 minutes at 4°C after adding 1 mL phosphate buffer saline (PBS) and was thereafter washed twice with PBS at 5000g for 15 minutes at 4°C. After washing, the biopsy samples were homogenised and resuspended in trizol. RNA isolation was then carried using trizol method as described by Abhishek et al, 2018 [27]. RNA was also isolated from pleural fluid samples after pelleting down the mycobacteria at 4500g for 20 minutes at room temperature using same protocol. As positive control, RNA from log phase *M.tb* H37Rv culture grown in Soutan's media using TRIzol as described by Abhishek et al [27]. The concentration and purity of isolated RNA was checked on nanodrop and RNA integrity of samples which were used for microarray was analysed using 2100 Bioanalyzer (Agilent).

M. tb whole genome microarray and data analysis

The transcriptome analysis of *M. tb* in TBP biopsy samples (n=4)/ *M.tb* H37Rv was carried out using whole genome microarray with 200ng input RNA using Agilent's microarray platform as described by Abhishek et al, 2018. Briefly the RNA was amplified and labeled with Cyanine 3-CTP (Cy3) followed by purification and quantitation of amplified cRNA. The samples having specific activity >15(pmol Cy3/μgcRNA), were further hybridized to custom designed microarray slides and scanned on the Agilent Sure Scan Microarray Scanner followed by image extraction and data analysis as described by Abhishek et al, 2018. Functional categorization of differentially expressed genes (DEGs) was performed according to TubercuList database. To

delineate the metabolic pathways specifically expressed in TBP patients, data obtained from *M.tb* whole genome microarray was further analyzed by using BioCyc database.(BioCyc.org)

qRT-PCR (Microarray validation)

The data attained from microarray was further confirmed by quantitative real time PCR. The Specific primers for *M.tb* genes were designed using Primer BLAST (supplementary table 1). Unamplified RNA (1µg-2µg) from biopsy samples of TBP patients and *in-vitro* grown *M.tb* H37Rv culture (Control) were DNase treated followed by cDNA synthesis using kit from Bio-Rad Laboratories as per manufacturer's instructions and used for qRT-PCR on Qiagen Rotor-Gene Q realtime PCR machine using iTaq Universal SYBR Green Supermix® dye and *M.tb* gene-specific primers. For each gene, cDNA samples (reverse transcriptase positive, RT+) along with the no enzyme control (NEC, reverse transcriptase negative) were run in duplicates. No enzyme control reactions were used to check for DNA contamination. The amplification plot (CT, cycle threshold) and melt curve for the product specificity were analyzed. Comparative CT method ($\Delta\Delta CT$) was used to determine relative gene expression of the target genes [27, 33] using 16srRNA as an endogenous control and *M.tb* RNA from *in-vitro* culture as a reference. Student's t test with $p < 0.05$ was used for statistical analysis.

Development and validation of qRT-PCR assay for the diagnosis of tuberculous pleuritis

The mycobacterial transcriptomic signatures identified from pleural biopsy samples were further used for the development and validation of a mycobacterial RNA based molecular assay for the diagnosis of TBP using pleural fluid samples. A computer-generated random number table for total 216 TBP suspects included in this study was used and 1/4th (n=53) of the study subjects with known clinical diagnosis (TBP and non TBP disease controls) were allocated to development cohort for the development of molecular assay. Remaining 3/4th (n=163) of the study subjects were assigned to validation cohort for the validation of developed molecular assay using blinded samples with no information on clinical diagnosis at the time of assay. Two of the top upregulated mycobacterial genes in pleural biopsy samples as identified by microarray data and validated through qRT-PCR were used for the development and validation of RNA based diagnostic assay using pleural fluid samples. Briefly, isolated *M.tb* RNA from the pleural fluid of all the patients was assayed for presence or absence of two selected genes using qRT-PCR as described above.

The presence or absence of *M.tb* RNA in the pleural fluid samples was analyzed and a sample was considered as positive for TB when mycobacterial gene specific RNA was detected in duplicate tubes with Ct value (cycle threshold) less than 35 along with transcript specific melt peak. The RT+ sample was considered as positive for the presence of *M.tb* RNA only if the RT+ and RT- samples differed in ≥ 1 Ct. [27].

RESULTS

In current study, total 261 TB pleuritis (TBP) suspects were recruited, out of which, 45 were excluded owing to indeterminate clinical diagnosis (n=24) or development of TB of any other site (n=7) or having both TB and malignancy simultaneously (n=4) or having other infectious diseases including HCV and COVID (n=4) or insufficient quantity of isolated RNA (n=6). Finally, 216 TBP suspects were included in the study. These study subjects (n=216) were divided into pleural TB (n=54) and non-TB controls (n=162) as per the composite reference standard (Figure1). Non-TB controls included patients with malignant pleural effusion (n=120) and other respiratory diseases (n=42) (Figure 1). Table 1 and 2 depict the demographic and clinical details of recruited study subjects.

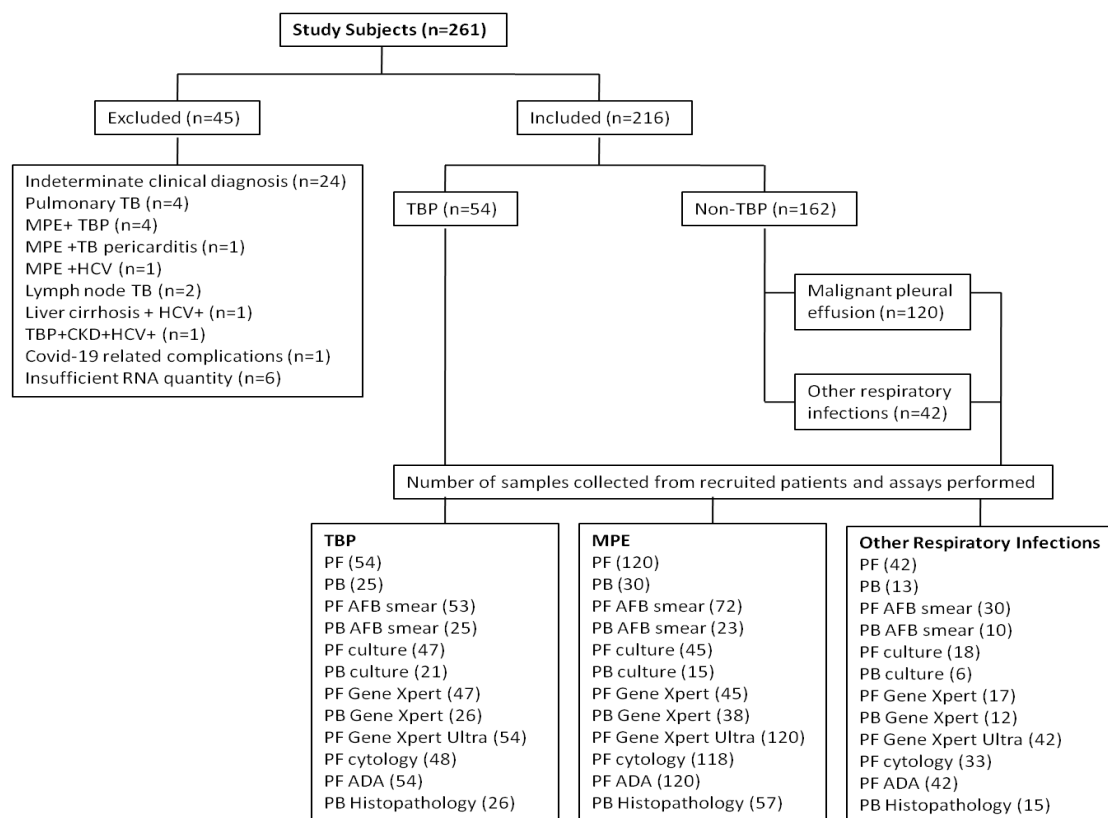


Figure 1: Consort diagram for the study showing details of enrolled study subjects. Values in parentheses show the number of patients in respective categories. ADA: Adenosine deaminase, AFB: Acid fast bacilli, PB: Pleural biopsy, PF: Pleural fluid, CKD: Chronic kidney disease, HCV: Hepatitis C virus, MPE: Malignant pleural effusion, TBP: TB pleuritis.

Table 1: Demographic profile of TBP and non TBP subjects.

Characteristics	TBP (n=54)	Non-TBP (n=162)
Age (yr)	39.5	55.5
Males, n (%)	39 (72.22%)	86 (53.08%)
HIV status		
Positive	0 (0%)	0 (0%)
Negative	18 (33.33 %)	33 (20.37%)
Status Unknown	36 (66.66%)	129 (79.62%)
Mantoux Status		
Positive	7 (12.96%)	5 (3.08%)
Negative	7 (12.96 %)	11 (6.79%)
Status Unknown	40 (74.07%)	146 (90.12%)
BCG status		
Yes	45 (83.33%)	143 (88.27 %)
No	9 (16.66 %)	11 (6.79%)
Status Unknown	0 (0%)	8 (4.93%)

TBP: TB pleuritis, HIV: Human immunodeficiency syndrome, BCG: bacille Calmette-Guerin

Table 2: Clinical characteristics of TBP and non TBP study subjects.

Characteristics	TBP (n=54)	Non-TBP (n=162)
Fever		
Yes	34 (62.96%)	60 (37.03%)
No	20 (37.03 %)	102 (62.96%)
Chest Pain		
Yes	32 (59.25%)	107 (66.04%)
No	22 (40.74 %)	55 (33.95%)
Shortness of Breath		
Yes	45 (83.33%)	142 (87.65%)
No	9 (16.66 %)	20 (12.34%)
Cough		
Yes (with Expectoration)	18 (33.33 %)	70 (43.20%)
Yes (without Expectoration)	27 (50%)	60 (37.03%)
No	9 (16.66%)	32 (19.75%)
Hemoptysis		
Yes	3 (5.55%)	15 (9.25%)
No	51 (94.44 %)	147 (90.74%)
Night Sweat		
Yes	26 (48.14%)	49 (30.24%)
No	28 (51.85 %)	112 (69.13%)
Status Unknown	0 (0%)	1 (0.61%)
Loss of appetite		
Yes	31 (57.40%)	90 (61.11%)
No	23 (42.59 %)	72 (44.44%)
Weight Loss		
Yes	41 (75.92%)	113 (69.75%)
No	13 (24.07 %)	49 (30.24%)
Previous history of TB		
Yes	4 (7.40%)	11 (6.79 %)
No	50 (92.59%)	151 (93.20%)
Not Known	0 (0%)	0 (0%)

Number within parentheses indicates percent of study subjects within TBP and non TBP patient groups. TBP: TB pleuritis.

Transcriptomic profile of *Mycobacterium tuberculosis* (*M.tb*) in pleural biopsy samples of TBP patients:

Whole genome microarray of *M.tb* was carried out in pleural biopsy samples from confirmed TBP patients (n=4) in comparison to in-vitro grown *M.tb* H37Rv culture. Overall whole genome microarray of mycobacteria in TBP patients led to identification of 1856 differentially expressed

genes (DEGs) including 1365 upregulated and 491 downregulated genes (Supplementary table 3& 4). (Figure 2b & c) The differentially expressed mycobacterial transcripts were functionally categorized as per the Tuberculist database [34] (Figure 2d). Though most of the functional categories were enriched equally among the upregulated and downregulated genes, however, lipid metabolism and PE-PPE family genes were enriched only among upregulated genes (Table 3). Pathway analysis using BioCyc database [35] led to identification of significantly enriched ($p < 0.05$) metabolic pathways among both the upregulated and downregulated categories. A significant enrichment among the upregulated genes was observed in pathways belonging to fatty acid, cholesterol and sugar/galactose degradation along with some other pathways as those belonging to pyrimidine biosynthesis, transcription regulation, DNA repair and metal ion transport (supplementary table 5). On the contrary a significant enrichment among downregulated genes was observed in pathways belonging to translation and protein secretion and localization. Genes belonging to stress responses like hypoxia, oxidative stress, nutrient stress, starvation and response to host defense pathways were also significantly enriched among the downregulated genes (supplementary table 6).

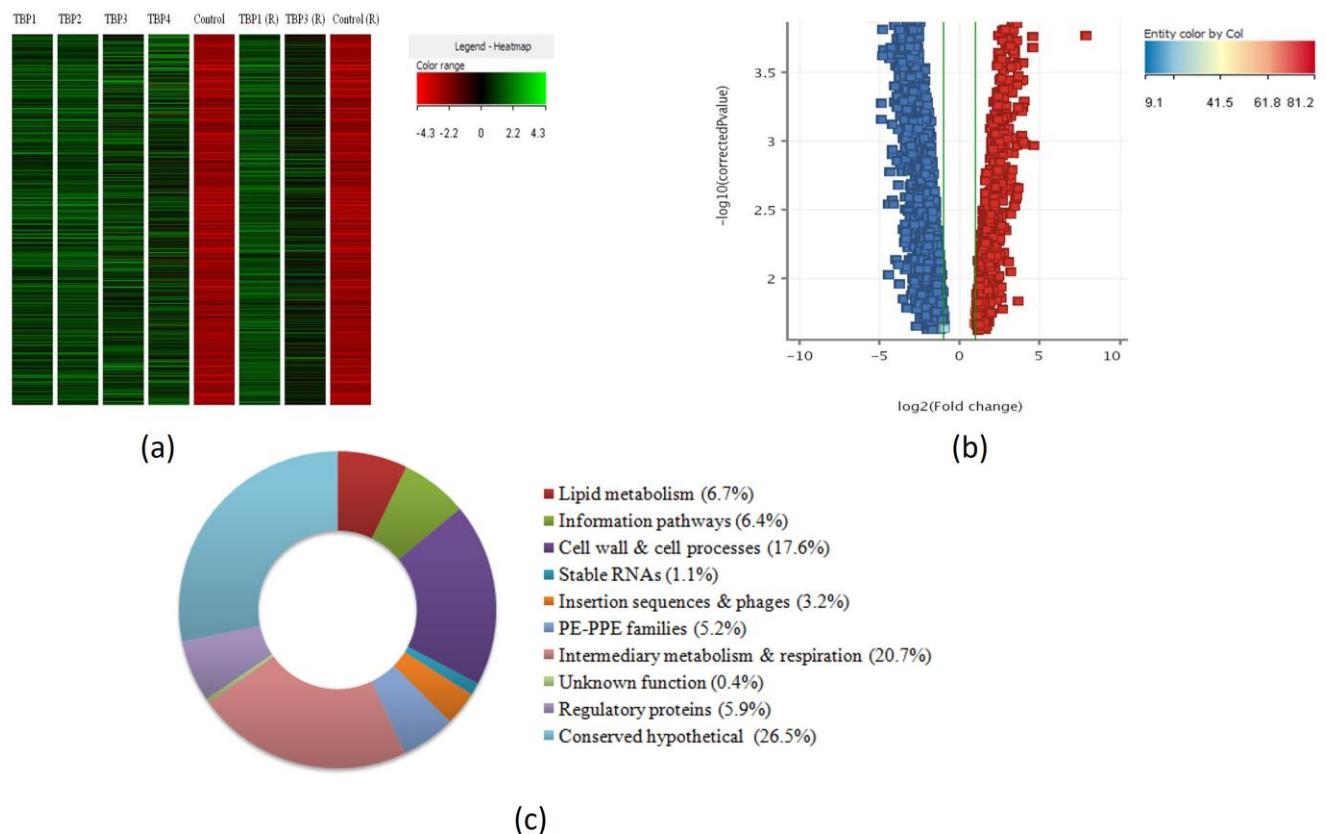


Figure 2: Transcriptome analysis of *M.tb* in biopsy samples of TBP patients through *M.tb* whole genome microarray. a) Heat map of mycobacterial transcriptome obtained in TBP and control samples (*in vitro* grown *M.tb* H37Rv culture) as obtained in whole genome microarray. Lane 1-4 (TBP samples),

Lane 5 (control; *in vitro* grown *M.tb* H37Rv culture), lane 6 (technical replicate of TBP sample 1), lane 7 (technical replicate of TBP sample 3) and lane 8 (technical replicate of control sample). b) Volcano plot of differentially expressed (upregulated (red) and down regulated (blue)) mycobacterial genes in biopsy samples from TBP patients as analyzed by GeneSpring software. Log₂ fold change is represented on x-axis and log₁₀ adjusted p-values on Y-axis. *M.tb*: *Mycobacterium tuberculosis*, TBP: TB pleuritis, RIN: RNA integrity number. c) The pie chart representing the percentage of differentially expressed *M.tb* genes in each functional category as per the Tuberculist database identified in the biopsy samples of TBP patients by *M.tb* whole genome microarray. TBP: TB pleuritis *M.tb*: *Mycobacterium tuberculosis*

Table 3: Functional categorization of differentially expressed mycobacterial genes (DEGs) in the biopsy samples of TBP patients.

Functional categories	Up regulated genes (Hypergeometric probability)	Down regulated genes (Hypergeometric probability)	Total
Virulence detoxification and adaptation (239)	76 (0.050)	32 (0.061)	108
Lipid metabolism (272)	96 (0.039)	29 (0.063)	125
Information pathways (242)	61 (0.001)	59 (1.33444E-08)	120
Cell wall & cell processes (772)	232 (0.003)	95 (0.045)	327
Stable RNAs (80)	8 (8.57412E-07)	14 (0.0405)	22
Insertion sequences & phages (147)	56 (0.030)	5 (0.0001)	61
PE-PPE families (168)	81 (1.37314E-05)	16 (0.063)	97
Intermediary metabolism & respiration (936)	289 (0.007)	96 (0.008)	385
Unknown function (15)	7 (0.113)	1 (0.302)	8
Regulatory proteins (198)	80 (0.005)	30 (0.031)	110
Conserved hypothetical (1042)	379 (0.001)	114 (0.001)	493
Total (4111)	1365	491	1856

Table depicts the functional categorization of upregulated and downregulated genes into various functional categories. Cut-off value of 2.0-fold change at p<0.05 was selected to consider differentially regulated expression of genes. Hypergeometric probability (value in parentheses) was applied to check for

statistically significant DEGs in each functional category and the statistically significant p value i.e. $p < 0.05$ is represented in bold numbers.

Microarray validation using qRT-PCR

Microarray data was validated using real time platform in biopsy samples with a subset of 10 genes including six upregulated genes (*Rv1586c*, *Rv2819c*, *Rv1971*, *Rv1582c*, *Rv2351c* and *Rv0974c*) and four downregulated genes (*Rv3872*, *Rv0983*, *Rv3875* and *Rv3871*) selected from the microarray analysis of TBP patients. Similar to the microarray data, the topmost upregulated genes *Rv1586c*, *Rv2819c* and *Rv1582c* showed an upregulated relative expression in real time RT-PCR with log2 fold change of 14.11, 13.64 and 12.43 respectively (Figure 3). However, in case of *Rv1971* and *Rv0974c*, no amplification was observed in any of the tested samples whereas in case of *Rv2351c*, amplification was seen in only one (Log2FC 0.92) sample. The downregulated genes *Rv3872*, *Rv3875*, *Rv0983* selected from microarray data showed downregulation in biopsy samples of TBP patients by qRT-PCR also (Figure 3) thus validating the microarray results. In case of *Rv3871*, no amplification and melt curve were observed in samples. Hence, the results of microarray data could be validated for 6/10 selected genes (Figure 3).

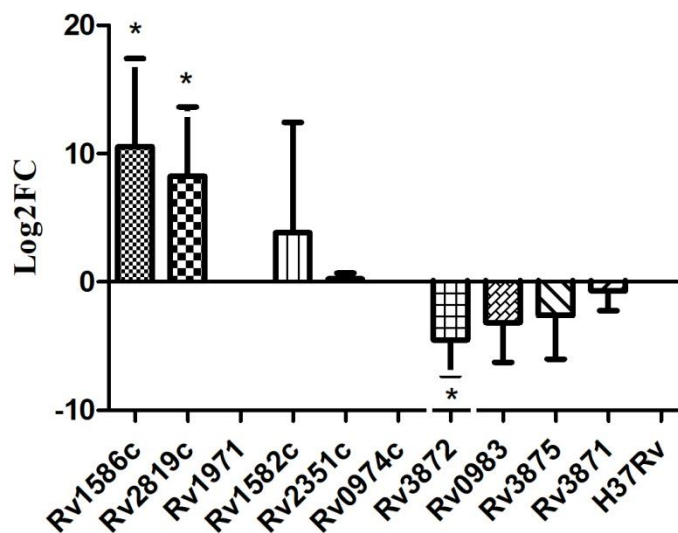


Figure 3: Validation of *M.tb* whole genome microarray by qRT-PCR. The relative expression of genes in both the upregulated (n=6) as well as downregulated (n=4) categories was analysed using RNA isolated from biopsy samples of TBP patients in comparison to *in-vitro* grown *M.tb* H37Rv. The bar graph represents log2 fold change in gene expression on y-axis, where positive values on y-axis indicate upregulation and negative values on y-axis indicate downregulation. For normalization, 16S rRNA was employed as reference gene (* $p < 0.05$ using student t-test for comparing the relative expression of each gene with *in-vitro* grown *M.tb* H37Rv). Each bar indicates log2 fold change in gene expression for the each of the gene obtained in 5-8 TBP samples with each sample analyzed in duplicates. *M.tb*: *Mycobacterium tuberculosis*, TBP: TB pleuritis. The relative expression of *Rv1586c* and *Rv2819c*

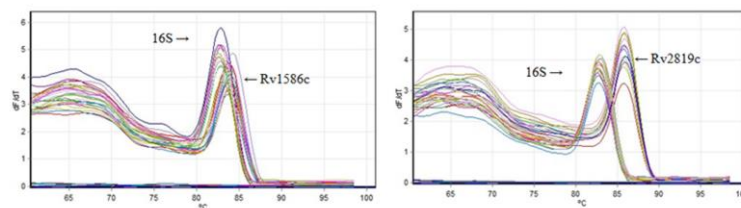
was analyzed in 8 TBP biopsy samples and the expression of *Rv1582c* was analyzed in 5 TBP biopsy samples.

Development and validation of RNA based molecular assay

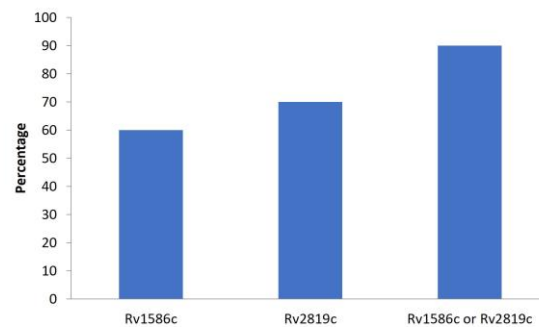
Development of molecular assay for TB pleuritis:

For developing real time RT-PCR based test using pleural fluid as sample, two topmost differentially upregulated *M.tb* genes *Rv1586c* (log2 FC: 9.89) and *Rv2819c* (FC: 8.09) whose expression were also validated through qRT-PCR were selected. To develop the assay, qualitative expression of selected genes i.e. *Rv1586c* and *Rv2819c* was first assessed in development cohort consisting of 10 TB pleuritis patients and 43 non-TB disease controls. Among the TBP patients, *Rv1586c* was detected in 6/10 patients thus showing sensitivity of 60%, however it also showed false detection in 3/43 disease controls thus giving a specificity of 93.02%. The other gene *Rv2819c* was detected in 7/10 TBP patients, showing a sensitivity of 70% and was falsely detected in 1/43 disease controls, thus having a specificity of 97.67 (Figure 4).

To increase the sensitivity of overall assay the data was further analyzed by looking at the presence of either of the genes for positive result. The analysis increased the sensitivity of detection to 90% with slight loss in the specificity which came out to be 93.02% (3/43). Thus, this analysis algorithm was taken forward for the evaluation of these genes in validation cohort (Figure 4b).



(a)



(b)

Figure 4: Development of qRT-PCR assay using mycobacterial genes *Rv1586c* and *Rv2819c* in pleural fluid samples of TBP and non-TBP subjects. a) Representative melt curve, b) The diagnostic performance of qRT-PCR. The data was analyzed by considering the positivity for *Rv1586c*, *Rv2819c* individually and either of the genes in each sample. Sample was considered positive based on Ct value (cut off 35) and presence of specific melt peaks for genes. Sensitivity=(TP)/(TP+FN), Specificity=(TN)/(FP+TN). Sen: Sensitivity, Spec: Specificity, TP: True positive, FN: False negative, TN: True negative, FP: False positive.

Validation of assay:

Rv1586c and *Rv2819c* based molecular assay was further validated in the blinded pleural fluid samples of validation cohort consisting of 163 TBP suspects. In the blinded analysis, using the algorithm from development cohort resulted in positivity in 43 samples with negative results in 120 samples (supplementary Figure 1).

After decoding, the true categorization of patients based on their final diagnosis using various laboratory and clinical criteria as described in methodology (Annexure III). Of 44 BP subjects, 35 samples were found as true positive and 9 subjects were falsely detected as negative by Amongst 119 non TBP subject, 111 samples were identified as negative and 8 were detected as falsely positive by qRT-PCR. Hence based on these results, the real time RT-PCR assay had a sensitivity of 79.55% (35/44 TBP) with specificity of 93.28% (8/119 non TBP patients) (Figure 5). Further on comparative evaluation of diagnostic potential of qRT-PCR based on these selected genes with available diagnostic modalities for TBP, real time RT-PCR assay showed better performance than either of existing diagnostic tests in pleural fluid (Supplementary table 8) except for histopathology that showed highest sensitivity and specificity of 94.74% and 100% respectively.

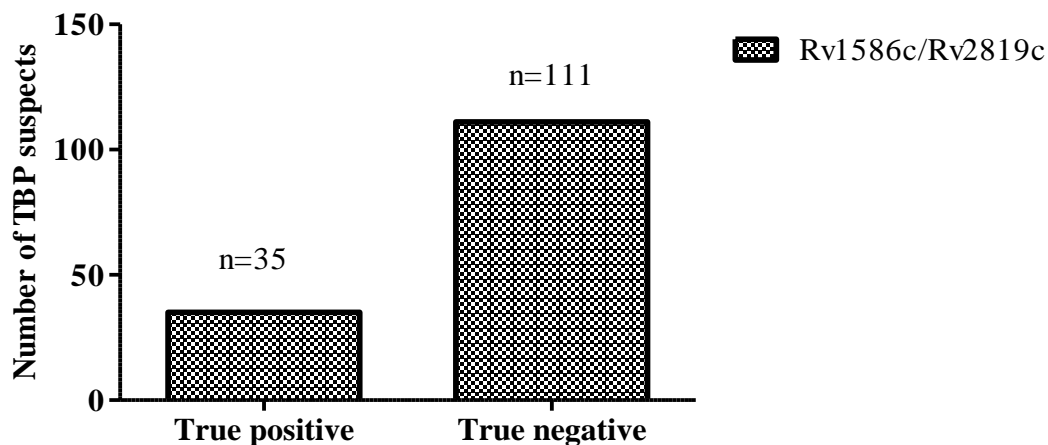


Figure 5: Performance of developed real time RT-PCR assay in pleural fluid samples of validation cohort :TBP suspects (n=163) were assessed for the presence of either of mycobacterial genes *Rv1586c/Rv2819c* and data was compared with the final diagnosis obtained for each patient using

composite reference standard to calculate the number of True positive (n=35), True Negative (n=111), False Positive (n=8) and False Negative (n=9) patients.

DISCUSSION

Routine diagnostic techniques utilizing pleural fluid suffer from limitations and do not provide reliable performance for TBP diagnosis. We hypothesized that the biomarkers which are expressed *in-vivo* at the site of infection during disease condition will serve as important diagnostic targets. Therefore, the mycobacterial transcriptome was studied in the biopsy samples obtained from confirmed TBP patients considering the fact that mycobacterial RNA isolated from pleural fluid might not be of optimum quantity and quality for microarray. Transcriptomic profile of *M.tb* in the biopsy samples from TBP patients demonstrated 1856 differentially expressed genes (DEGs) as compared to *in-vitro* grown *M.tb* H37Rv. The number of upregulated genes was higher in comparison to downregulated genes, a pattern similar to that observed in *M.tb* infected alveolar epithelial cells and macrophages but contrary to sputum samples of PTB patients [25,28,36,37]. The *in-vivo* upregulated mycobacterial transcripts in pleural biopsy samples were further used for the development of RNA based molecular assay in pleural fluid samples.

Several studies have previously used host transcriptomic signatures for TB diagnosis [38-40]. However, pathogen specific transcriptomic signatures seem to be more important for the specificity of an assay. For exploring the role of mycobacterial transcripts for the development of a molecular assay for pleural TB, microarray results were first validated on real time quantitative reverse transcription PCR (qRT-PCR) using unamplified RNA isolated from pleural biopsy samples obtained from TBP patients. Six out of 10 genes assessed by qRT-PCR showed same expression pattern as that observed in microarray. Some level of disagreement between microarray and qRT-PCR has also been observed in previous studies [26, 29] which may be due to multiple factors like the handling and storage conditions of samples, differences in the data interpretation strategies etc. [41] We selected two of the top upregulated genes validated through qRT-PCR i.e., *Rv1586* and *Rv2819c* belonging to different functional categories for development of an RNA based molecular assay using the pleural fluid samples to discriminate TBP and non TBP patients. *Rv1586c* (Probable PhiRv1 integrase) belongs to the category of Probable PhiRv1 phage proteins of prophage-like elements ϕ Rv1 of *M.tb* and encodes putative integrase of the serine recombinase type. The prophage-like elements ϕ Rv1 is specifically expressed in *M.tb* and absent in BCG. The second selected gene *Rv2819c*, encodes for a hypothetical protein (375 amino acid) and not much is known about it.

The qRT-PCR assay based on these two genes *Rv1586c* and *Rv2819c* was initially developed in study subjects comprising development cohort and having known diagnosis on the basis of routine laboratory and clinical diagnostics, however blinded samples were used for validation part of the study. The data of qRT-PCR was analysed using the MIQE guidelines which have

previously been used by several studies reporting qRT-PCR data including the diagnostic assays developed for COVID testing based on qRT-PCR [42-44]. Using these criteria, when positivity of only *Rv1586c* or *Rv2819c* was considered, it led to diagnosis of TBP patients with sensitivity and specificity of 60-70% and 93-97% respectively. Analysis of data indicated that although many of the samples showed positivity for both the genes, but several samples also showed the positivity to either of these two genes. Hence, to further improve the sensitivity of the assay, positivity of either of genes for a positive qRT-PCR test was considered which resulted in improved sensitivity without significant compromise in specificity for differentiating TBP patients from non TBP study subjects. Hence for validation cohort, results were analyzed considering the positivity for either of the genes (*Rv1586c* / *Rv2819c*) in the same pleural fluid sample. Decoding the qRT-PCR results obtained in TBP suspects of validation cohort based on the CRS indicated 79.6% sensitivity and 93.02% specificity of the assay for TBP diagnosis. In comparison to the RNA based assay developed in this study, the sensitivity of Xpert MTB/RIF and Xpert MTB/RIF Ultra was lower (Supplementary Table 6). Previous studies have also shown low sensitivity of Xpert MTB/RIF in pleural fluid [45-47]. although recent studies on Xpert MTB/RIF Ultra have shown improved performance over Xpert MTB/RIF [24,48,49]. The test developed in the current study although has a better sensitivity than ultra, however specificity is slightly lower. Moreover, the test being an RNA based assay could be advantageous over DNA based Xpert MTB/RIF and Ultra that cannot discriminate between live and dead bacilli leading to false positivity in previously treated individuals [25, 50]. But assay developed in current study cannot be used for detection of antibiotic resistance as in case of Xpert MTB/RIF and Xpert MTB/RIF Ultra. Though RNA based assay will not be a user-friendly assay and will require trained personnel, but the widespread use of a similar assay for diagnosing SARS-COV2 during the COVID pandemic has provided mankind with enough experience to handle such assays even at hospitals other than tertiary care centers.

As a secondary objective, the in vivo expressed mycobacterial transcriptome data was also analyzed in depth to understand the pathogenesis of pleural TB. The differentially expressed upregulated and downregulated genes were categorized into different functional categories as per tuberculist database [34]. Genes of PE-PPE family and lipid metabolism were significantly enriched in the upregulated pathways (Table 3). Number of PE-PPE family proteins are implicated in *M.tb* pathogenesis via promoting the attachment of microbe to the host cells [51,52]. A significant enrichment of fatty acid degradation, utilization and assimilation along with cholesterol degradation pathways was observed among the upregulated category. Fatty acids and lipids including cholesterol are vital components of mycobacterial cell wall and also essential for carbon and energy sources for *M.tb* adaptation in host microenvironment to establish and maintain chronic TB infection [53-57]. Pyrimidine biosynthesis, transcription regulation and DNA repair pathways known to play an important role during *M.tb* infection and disease progression were also upregulated [58, 59]. Several down regulated pathways were also significantly enriched and primarily included translation, protein secretion and localization similar to that reported previously in sputum mycobacterial cells [26,37] and could be considered

as an adaptive strategy in host environment [60]. Overall, it seems that mycobacteria in pleural space is engaged in energy generation using fatty acid degradation and is increasing its virulent properties with the upregulation of PE-PPE genes. These genes along with DNA repair genes are also helping the bacteria to deal with the host immune system and various stresses.

In conclusion, the transcriptome analysis of *Mycobacterium tuberculosis* in biopsy samples of TB pleuritis patients led to newer insight into the in-vivo expression of mycobacterial transcriptomic signatures in the pleural space during TBP. The selected pathogenic RNA signatures from the M. tb transcriptome in TBP which are highly expressed in-vivo have shown promising results for the development of RNA based diagnostic real time PCR assay using pleural fluid. However, it needs further optimization to develop it into user friendly assay devoid of technical complexities of real time RT-PCR.

Funding: The current research was funded by Department of Biotechnology, New Delhi, India vide Grant No. TB.BT/PR20863/MED/30/1889/2017.

Data Availability: The gene expression data obtained in the current study has been submitted to Gene Expression Omnibus (GEO) vide accession no. GSE231472.

Conflict of Interest: None

REFERENCES

1. WHO. Global TB report 2021. Geneva, Switzerland.
2. Cascio, C., Kaul, V., Dhooria, S., Agrawal, A. & Chaddha, U. Diagnosis of tuberculous pleural effusions: A review. *Respir. Med.* 2021; 188: 106607.
3. Light, R.W., Bays, A.M., Pierson, D.J., Zhai, K., Lu, Y. & Shi, H. Tuberculous pleural effusion. *J. Thorac. Dis.* 2012; 57: 1–9.
4. Bielsa, S., Acosta, C., Pardina, M., Civit, C. & Porcel JM. Tuberculous pleural effusion: Clinical characteristics of 320 patients. *Arch Bronconeumol (Engl Ed)*. 2019; 55: 17–22.
5. Bueno, E., Clemente, M., Castro, B., Martín, L., Ramos, S., Panizo, A., et al. Cytologic and bacteriologic analysis of fluid and pleural biopsy specimens with cope's needle: Study of 414 patients. *Arch. Intern. Med.* 1990; 150: 1190–1194.
6. Seibert, A., Haynes, J., Middleton, R. & Bass, J. Tuberculous pleural effusion: Twenty-year experience. *Chest.* 1991; 99: 883–886
7. Vorster, M., Allwood, B., Diacon, A. & Koegelenberg, C. Tuberculous pleural effusions: advances and controversies. *J. Thorac. Dis.* 2015; 7: 981–991.
8. Berger, H. & Mejia, E. Tuberculous pleurisy. *Chest.* 1973; 63: 88–92.

9. Epstein, D., Kline, L., Albelda, S. & Miller, W. Tuberculous pleural effusions. *Chest*. 1987; 91: 106–109.
10. Gopi, A., Madhavan, S., Sharma, S. & Sahn, S. Diagnosis and treatment of tuberculous pleural effusion in 2006. *Chest*. 2007; 131: 880–889
11. Wang Z, Xu, L., Wu, Y., Wang, X., Yang, Y., Zhang, J., et al. Diagnostic value and safety of medical thoracoscopy in tuberculous pleural effusion. *Respir Med*. 2015; 109: 1188–1192.
12. Ruan, S., Chuang, Y., Wang, J., Lin, J., Chien, J., Huang, C., et al. Revisiting tuberculous pleurisy: pleural fluid characteristics and diagnostic yield of mycobacterial culture in an endemic area. *Thorax*. 2012; 67: 822–827.
13. Diacon, A., Wal, B. Van de, Wyser, C., Smedema, J., Bezuidenhout, J., Bolliger, C., et al. Diagnostic tools in tuberculous pleurisy: a direct comparative study. *Eur Respir J*. 2003; 22: 589–591.
14. Bibby, A. & Maskell, N. Pleural biopsies in undiagnosed pleural effusions; Abrams vs image-guided vs thoracoscopic biopsies. *Curr Opin Pulm Med*. 2016; 22: 392–398.
15. Porcel, J. Tuberculous pleural effusion. *Lung*. 2009; 187: 263–270.
16. Light, R. Useful tests on the pleural fluid in the management of patients with pleural effusions. *Curr Opin Pulm Med*. 1999; 5: 245–249.
17. Sharma, S. & Mohan, A. Adenosine deaminase in the diagnosis of tuberculosis pleural effusion. *Indian J Chest Dis Allied Sci*. 1996; 38: 69–71.
18. Trajman, A. & Luiz, R. McNemar chi² test revisited: comparing sensitivity and specificity of diagnostic examinations. *Scand J Clin Lab Invest*. 2008; 68: 77–80.
19. Trajman, A., Pai, M., Dheda, K., Van Zyl Smit, R., Zwerling, A., Joshi, R., et al. Novel tests for diagnosing tuberculous pleural effusion: What works and what does not? *Eur. Respir. J*. 2008; 31: 1098–1106.
20. Aggarwal, A.N., Agarwal, R., Dhooria, S., Prasad, K.T., Sehgal, I.S. & Muthu, V. Unstimulated pleural fluid interferon gamma for diagnosis of tuberculous pleural effusion: A systematic review and meta-analysis. *J. Clin. Microbiol*. 2021; 59: 2112–2132.
21. Meldau, R., Randall, P., Pooran, A., Limberis, J., Makambwa, E., Dhansay, M., et al. Same-Day Tools, Including Xpert Ultra and IRISA-TB, for rapid diagnosis of pleural tuberculosis: a prospective observational study. *J. Clin. Microbiol*. 2019; 57: e00614-19.
22. Wu, X., Tan, G., Gao, R., Yao, L., Bi, D., Guo, Y., et al. Assessment of the Xpert MTB/RIF Ultra assay on rapid diagnosis of extrapulmonary tuberculosis. *Int J Infect Dis*. 2019; 81: 91–96.

23. Wang, G., Wang, S., Yang, X., Sun, Q., Jiang, G., Huang, M., et al. Accuracy of Xpert MTB/RIF Ultra for the diagnosis of pleural TB in a multicenter cohort study. *Chest*. 2020; 157: 268–275.
24. Gao, S., Wang, C., Yu, X., Teng, T., Shang, Y., Jia, J., et al. Xpert MTB/RIF Ultra enhanced tuberculous pleurisy diagnosis for patients with unexplained exudative pleural effusion who underwent a pleural biopsy via thoracoscopy: A prospective cohort study. *Int J Infect Dis*. 2021; 106: 370–375.
25. Theron, G., Venter, R., Calligaro, G., Smith, L., Limberis, J., Meldau, R., et al. b. Xpert MTB/RIF Results in Patients With Previous Tuberculosis: Can We Distinguish True From False Positive Results? *Clin. infect dis*. 2016; 62: 995–1001.
26. Sharma, S., Ryndak, M.B., Aggarwal, A.N., Yadav, R., Sethi, S., Masih, S., et al. Transcriptome analysis of mycobacteria in sputum samples of pulmonary tuberculosis patients. *PLoS One*. 2017; 12: e0173508.
27. Abhishek, S., Saikia, U., Gupta, A., Bansal, R., Gupta, V., Singh, N., et al. Transcriptional profile of mycobacterium tuberculosis in an in vitro model of intraocular tuberculosis. *Front Cell Infect Microbiol*. 2018; 8: 1–20.
28. Fontán, P., Aris, V., Ghanny, S., Soteropoulos, P. & Smith, I. Global transcriptional profile of *Mycobacterium tuberculosis* during THP-1 human macrophage infection. *Infect Immun*. 2008; 76: 717–725.
29. Ryndak, M.B., Singh, K.K., Peng, Z., Zolla-Pazner, S., Li, H., Meng, L., et al. b. Transcriptional profiling of *Mycobacterium tuberculosis* replicating ex vivo in blood from HIV- and HIV+ subjects. *PLoS One*. 2014; 9: e94939.
30. Schnappinger, D., Sabine, E., Voskuil, M., Liu, Y., Mangan, J., Monahan, I., et al. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J Exp Med*. 2003; 198: 693–704.
31. Ward, S., Abomoelak, B., Marcus, S. & Talaat, A. Transcriptional profiling of mycobacterium tuberculosis during infection: lessons learned. *Front Microbiol*. 2010; 1: 121.
32. Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., et al. Unique Transcriptome Signature of *Mycobacterium tuberculosis* in Pulmonary Tuberculosis †. *Infect Immun*. 2006; 74: 1233–1242.
33. Livak, K. & Schmittgen, T. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(-Delta Delta C(T)) Method. *Methods*. 2001; 25: 402–408.
34. Lew, J., Kapopoulou, A., Jones, L. & Cole, S. TubercuList--10 years after. *Tuberculosis (Edinb)*. 2011; 91: 1–7.

35. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A et al. *Brief Bioinform.* 2019; 20:1085-1093.
36. Ryndak, M., Singh, K., Peng, Z. & Laal, S. a. Transcriptional profile of mycobacterium tuberculosis replicating in type II alveolar epithelial cells. *PLoS One.* 2015; 10: e0123745. Public Library of Science.
37. Garton NJ, Waddell SJ, Sherratt AL, Lee SM, Smith RJ, Senner C et al Cytological and transcript analyses reveal fat and lazy persistor-like bacilli in tuberculous sputum. *PLoS Med.* 2008;5:e75.
38. Warsinske, H., Vashisht, R. & Khatri, P. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. *PLoS Med.* 2019; 16.
39. Hoang, L., Jain, P., Pillay, T., Tolosa-Wright, M., Niazi, U., Takwoingi, Y., et al. Transcriptomic signatures for diagnosing tuberculosis in clinical practice: a prospective, multicentre cohort study. *Lancet Infect Dis.* 2021; 21: 366–375.
40. Kanabalan, R., Lee, L., Lee, T., Chong, P., Hassan, L., Ismail, R., et al. Human tuberculosis and Mycobacterium tuberculosis complex: A review on genetic diversity, pathogenesis and omics approaches in host biomarkers discovery. *Microbiol. Res.* 2021; 246: 126674.
41. Morey, J., Ryan, J. & Van Dolah, F. Microarray validation: Factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online.* 2006; 8: 175–193.
42. Calura, E., Paracchini, L., Fruscio, R., DiFeo, A., Ravaggi, A., Peronne, J., et al. A prognostic regulatory pathway in stage I epithelial ovarian cancer: new hints for the poor prognosis assessment. *Ann Oncol.* 2016; 27: 1511–1519.
43. Bayés-Genis, A., Lanfear, D., de Ronde, M., Lupón, J., Leenders, J., Liu, Z., et al. Prognostic value of circulating microRNAs on heart failure-related morbidity and mortality in two large diverse cohorts of general heart failure patients. *Eur. J. Hear. Fail.* 2018; 20: 67–75.
44. Bustin, S., Coward, A., Sadler, G., Teare, L. & Nolan, T. CoV2-ID, a MIQE-compliant sub-20-min 5-plex RT-PCR assay targeting SARS-CoV-2 for the diagnosis of COVID-19. *Sci. Rep. reports.* 2020; 10.
45. Lusiba, J., Nakiyingi, L., Kirenga, B., Kiragga, A., Lukande, R., Nsereko, M., et al. Evaluation of Cepheid's Xpert MTB/Rif test on pleural fluid in the diagnosis of pleural tuberculosis in a high prevalence HIV/TB setting. *PLoS One.* 2014; 9: e102702.
46. Li, C., Liu, C., Sun, B., Zhang, W., Wang, Y., Sun, J., et al. Performance of Xpert® MTB/RIF in diagnosing tuberculous pleuritis using thoracoscopic pleural biopsy. *BMC Infect Dis.* 2020; 20: 840.

47. Denkinger, C., Schumacher, S., Boehme, C., Dendukuri, N., Pai, M. & Steingart, K. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. *Eur Respir J.* 2014; 44: 435–446.
48. Osei Sekyere, J., Maphalala, N., Malinga, L., Mbelle, N. & Maningi, N. A Comparative Evaluation of the New Genexpert MTB/RIF Ultra and other Rapid Diagnostic Assays for Detecting Tuberculosis in Pulmonary and Extra Pulmonary Specimens. *Sci. Rep.* 2019; 9: 1–9.
49. Aggarwal AN, Agarwal R, Dhooria S, Prasad KT, Sehgal IS, Muthu V. Xpert MTB/RIF Ultra versus Xpert MTB/RIF for diagnosis of tuberculous pleural effusion: A systematic review and comparative meta-analysis. *PLoS One.* 11;17:e0268483.
50. Arend SM, van Soolingen D. Performance of Xpert MTB/RIF Ultra: a matter of dead or alive. *Lancet Infect Dis.* 2018;18:8-10.
51. Brennan, M.J. The enigmatic PE/PPE multigene family of mycobacteria and tuberculosis vaccination. *Infect Immun.* 2017; 85: e00969-16.
52. Akhter, Y., Ehebauer, M., Mukhopadhyay, S. & Hasnain, S. The PE/PPE multigene family codes for virulence factors and is a possible source of mycobacterial antigenic variation: Perhaps more? *Biochimie.* 2012; 94: 110–116.
53. Boshoff, H. & Barry, C. Tuberculosis - Metabolism and respiration in the absence of growth. *Nat. Rev. Microbiol.* 2005; 3: 70–80.
54. Muñoz-Elías, E. & McKinney, J. Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. *Nat. Med.* 2005; 11: 638–644.
55. Gago, G., Diacovich, L. & Gramajo, H. Lipid metabolism and its implication in mycobacteria-host interaction. *Curr. Opin. Microbiol.* 2018; 41: 36–42.
56. Wilburn, K., Fieweger, R. & VanderVen, B. Cholesterol and fatty acids grease the wheels of Mycobacterium tuberculosis pathogenesis. *Pathog. Dis.* 2018; 76: fty021.
57. Pawełczyk, J., Brzostek, A., Minias, A., Płociński, P., Rumijowska-Galewicz, A., Strapagiel, D., et al. Cholesterol-dependent transcriptome remodeling reveals new insight into the contribution of cholesterol to Mycobacterium tuberculosis pathogenesis. *Sci. Rep.* 123AD; 11: 12396.
58. Kantardjieff, K., Vasquez, C., Castro, P., Warfel, N., Rho, B., Lekin, T., et al. Structure of pyrR (Rv1379) from Mycobacterium tuberculosis: a persistence gene and protein drug target. *Acta Crystallogr D Biol Crystallogr.* 2005; 61: 355–364.
59. Warner, D. Mycobacterium tuberculosis metabolism. *Cold Spring Harb Perspect Med.* 2014; 5: a021121.

60. Rosenberg A, Sinai L, Smith Y, Ben-Yehuda S. Dynamic expression of the translational machinery during *Bacillus subtilis* life cycle at a single cell level. PLoS One. 2012;7:e41921.