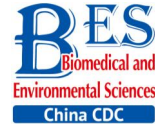# Original Article

Screening of Serum Biomarkers for Distinguishing between Latent and Active Tuberculosis Using Proteome Microarray[*]

CAO Shu Hui[1,2], CHEN Yan Qing[1], SUN Yong[3], LIU Yang[4], ZHENG Su Hua[4],

ZHANG Zhi Guo[5], and LI Chuan You[1,#]

1. Department of Bacteriology and Immunology, Beijing Key Laboratory on Drug-Resistant Tuberculosis Research, Beijing Tuberculosis and Thoracic Tumor Research Institute/Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 2. Department of Laboratory Center, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China; 3. Department of Clinical Laboratory, Beijing Tuberculosis and Thoracic Tumour Research Institute/Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 4. Department of Epidemiological Research, Beijing Tuberculosis and Thoracic Tumour Research Institute/Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 5. Institute of Tuberculosis Prevention and Control of District Changping, Beijing 102200, China

Abstract

Objective   To identify potential serum biomarkers for distinguishing between latent tuberculosis infection (LTBI) and active tuberculosis (TB).

Methods   A proteome microarray containing 4,262 antigens was used for screening serum biomarkers of 40 serum samples from patients with LTBI and active TB at the systems level. The interaction network and functional classification of differentially expressed antigens were analyzed using STRING 10.0 and the TB database, respectively. Enzyme-linked immunosorbent assays (ELISA) were used to validate candidate antigens further using 279 samples. The diagnostic performances of candidate antigens were evaluated by receiver operating characteristic curve (ROC) analysis. Both antigen combination and logistic regression analysis were used to improve diagnostic ability.

Results   Microarray results showed that levels of 152 *Mycobacterium tuberculosis* (Mtb)-antigen-specific IgG were significantly higher in active TB patients than in LTBI patients ($P < 0.05$), and these differentially expressed antigens showed stronger associations with each other and were involved in various biological processes. Eleven candidate antigens were further validated using ELISA and showed consistent results in microarray analysis. ROC analysis showed that antigens Rv2031c, Rv1408, and Rv2421c had higher areas under the curve (AUCs) of 0.8520, 0.8152, and 0.7970, respectively. In addition, both antigen combination and logistic regression analysis improved the diagnostic ability.

Conclusion   Several antigens have the potential to serve as serum biomarkers for discrimination between LTBI and active TB.

Key words: LTBI; Active TB; Proteome microarray; Serum biomarkers

INTRODUCTION

Tuberculosis (TB) is a major cause of morbidity and mortality due to infectious diseases worldwide. According to data from the World Health Organization, approximately 10.4 million new cases of TB occurred, and 1.8 million people died from TB worldwide in 2015[1]. Among individuals infected with *Mycobacterium tuberculosis* (Mtb), approximately 5% develop active TB over 2-5 years, whereas the remaining 95% develop latent TB infection (LTBI) without clinical, radiological, or bacteriological evidence of active TB, but show immune responses to Mtb infection[2-4]. Currently, nearly one-third of the world's population have LTBI. However, these individuals have a 10% chance of reactivation of the latent infection, which may ultimately progress to active TB during their lifetimes. So, individuals with LTBI are potential reservoirs of active TB[1,5]. Therefore, identification of LTBI is urgent for decreasing the risk of developing active TB, especially in certain high-risk populations. At the same time, early diagnosis followed by treatment of active TB is the most effective method for controlling TB epidemics.

Currently, early detection of individuals with Mtb infection mainly relies on the host's cell-mediated immune response to pathogen-specific antigens. Examples of such tests are the *in vivo* tuberculin skin test (TST) and *in vitro* interferon gamma release assays (IGRAs)[6]. The TST is based on the use of a purified protein derivative (PPD), which is isolated from culture filtrates of the tubercle bacilli. Specific antigens in the PPD can elicit a cross-reaction with other mycobacterium species, including the BCG vaccine, thereby showing poor specificity[7,8]. IGRAs are based on the measurement of interferon gamma (IFN-γ) released from T-lymphocytes exposed to Mtb antigens, which is easier, faster, and more sensitive than the TST. Hence, IGRAs are commonly used by clinical physicians[9]. In particular, IGRAs can differentiate between Mtb infection and BCG vaccination and exclude interference of most non-tuberculosis mycobacterium. However, neither of these tests distinguish between active TB and LTBI[10,11]. Thus, there is an urgent requirement for rapid and simple tests that indicate disease status.

Serological assays have a long history and have been widely used for the diagnosis of several infectious diseases because of convenient sampling, low cost, easy procedure, and short turnaround time.

Currently, several serological commercial kits based on antibody detection for TB identification have been developed, such as Linonex TB kits, InBios Active TbDetect IgG ELISA, IBL *M. tuberculosis* IgG ELISA, and Anda Biologicals TB ELISA[12,13]; however, these kits do not yield accurate results, which limits their clinical use. Since serological detection in a simple dipstick format can incorporate various antigens, investigation of serum biomarkers for both diagnosing TB and determining disease status is a worthwhile endeavor.

The present study aimed to screen potential serum biomarkers for distinguishing between LTBI and active TB at the system level using the Mtb proteome microarray containing 4,262 antigens. We observed that the levels of 152 Mtb antigen-specific IgG antibodies were higher in the active TB group than in the LTBI group, and these antigens showed both stronger associations with each other and involvement in various biological processes. We further validated 11 candidate antigens with ELISA and used ROC analysis to evaluate the diagnostic performance of these antigens. In addition, both antigen combination and the logistic regression model exhibited better ability than any single antigen for distinguishing between LTBI and active TB patients, which provides the necessary groundwork required for establishing a new method for determining TB disease status.

MATERIALS AND METHODS

Study Population

This study was approved by the ethics committee of the Beijing Chest Hospital, Capital Medical University (Beijing, China) in accordance with the Declaration of Helsinki. Written informed consent for the use of samples was obtained from all participants.

From June 2014 to December 2015, 319 subjects were included in this study and divided into three groups: active TB patients group (active TB group), LTBI subjects group (LTBI group), and healthy control group (HC group). Active TB patients were recruited from the Beijing Chest Hospital; LTBI subjects and healthy controls were recruited from the Institute of Tuberculosis Prevention and Control of Changping district (Beijing, China). Active TB patients were diagnosed according to the guidelines of 'Pulmonary TB Diagnosis and Therapy' authored by the Tuberculosis Branch Association of the Chinese

Medical Association, which included abnormal chest radiography, positive Ziehl-Neelsen-stained sputum smear and/or sputum culture, and the presence of TB clinical symptoms. All active TB patients had been treated for no longer than 2 weeks at the time of their blood sampling. Every LTBI subject showed a positive response to both the TST (> 10 mm) and the T-spot TB assay, whereas healthy controls showed negative responses to both the TST (< 5 mm) and the T-spot TB assay. Both LTBI subjects and healthy controls were free of all clinical symptoms or abnormal chest radiographic findings, which are indicative of active TB. In addition, all subjects with HIV and those treated with immunosuppressive medications were excluded. Assays for TST, anti-HIV, and T-spot were performed using TB-PPD (Sanroadbio, Beijing, China), ELISA reagent kit (Livzon Diagnostic Inc., Zhuhai, China), and T-spot TB reagent kit (Oxford Immunotec, Abingdon, UK), respectively, according to manufacturers' instructions.

Serum Samples

Three milliliters peripheral venous blood was collected. Sera were obtained by centrifugation at 1,509 ×*g* for 10 min. Then, 400 µL aliquots of sera were aliquoted into 0.5 mL clean eppendorf tubes and stored at -80 °C until further use.

Serum Profiling on Mtb Proteome Microarray

The Mtb proteome microarrays used in our study were purchased from BCBIO (Guangzhou, China). Microarrays comprise 3,829 proteins encoded by genes of H37Rv (Mtb standard strain) and 433 proteins encoded by genes of CDC1551 (pathogenic strain), which were spotted in duplicates on polymer slides (polymer-slide H, CapitalBio)[14]. Expression of these GST-tagged recombinant proteins was verified in a *Saccharomyces cerevisiae* expression system. In addition, positive controls of human IgG and IgM, and negative controls of bovine serum albumin (BSA) were also printed in duplicates on the microarrays.

First, microarrays were blocked for 1 h at room temperature with agitation in blocking buffer [3% BSA in 1× TBST (Tris-buffered saline plus 0.1% Tween 20 detergent, pH 7.4)]. Three milliliters of serum samples (1:50 dilution in TBST) were overlaid onto protein microarrays and then incubated at room temperature for 3 h. After washing three times for 5 min each with TBST, microarrays were probed with goat anti-human IgG conjugated with Cy3 (Jackson Laboratory, PA, USA) diluted 1:1,000 in TBST and incubated in a dark room at room temperature for 45 min. Next, microarrays were washed three times with TBST and then twice with double-distilled water in the dark. Finally, microarrays were dried in a SlideWasher (CapitalBio, Beijing, China) at room temperature, and scanned at 532 nm in a GenePix 4200A (Molecular Devices, CA, USA). Fluorescence data were analyzed using GenePix Pro 6.0 software (Molecular Devices, CA, USA).

ELISA

Mtb antigens were purchased from BCBIO (Guangzhou, China). Each well of the 96-well flat-bottom plates (Thermo, Copenhagen, Denmark) was coated with 100 µL of 5 µg/mL individual antigens in coating buffer (0.1 mol/L carbonate/bicarbonate, pH 9.6) and stored at 4 °C overnight. Plates were washed three times with phosphate-buffered saline containing 0.05% Tween 20 (PBST) for 5 min each, and then blocked with 200 µL/well PBST containing 1% bovine serum albumin (PBST-B) at room temperature for 2 h in a humidified chamber. After washing three times with PBST, 100 µL of serum samples diluted 1:400 in PBST-B were added to antigen-coated wells and incubated at room temperature for 1 h. Subsequently, plates were washed five times, 100 µL/well of horseradish peroxidase-conjugated anti-human IgG antibody (CWBiotech, Beijing, China) diluted 1:30,000 in PBST-B was added, and incubated at room temperature for 1 h. Next, plates were washed five times, and developed using 100 µL/well of TMB substrate (BD, NJ, USA) in a dark room for 10 to 15 min until visible color was apparent; the reaction was stopped using 50 µL/well of 2 mol/L sulfuric acid. Finally, optical density was determined at 450 nm using an automatic microplate reader (Perlong, Beijing, China).

Analysis of Protein Microarray Data

Protein microarray data were obtained using GenePix Pro 6.0. The background signal of the raw data was corrected to eliminate variations between arrays, followed by normalization using the limma package of the R programming language (http://www.protein-microarray.com). This pre-processed data was used for further analysis. Differences in microarray data between active TB and LTBI were analyzed with a *t*-test for two independent samples. Fold change was defined as the log-transformed ratio of active TB to LTBI. Hierarchical cluster, obtained using R statistical software, was

represented by log-transformed values. Protein-protein interaction networks were analyzed using STRING 10.0 (http://string-db.org/) and visualized with a confidence level of 0.25 as the parameter setting.

Statistical Analysis

Differences in age and gender among the three groups were analyzed using the Kruskal-Wallis test and Chi-square test, respectively. Differences in optical density values among the three groups were analyzed using a one-way analysis of variance with Bonferroni comparisons. Differences between paired proportions were analyzed with McNemar's test. P values less than 0.05 were considered statistically significant.

For each antigen, diagnostic performance was examined using ROC curve analysis to determine AUCs and their 95% confidence intervals (CI). The cutoff level of each antigen was determined at the maximum Youden's index (YI = Sensitivity + Specificity − 1).

For analyzing antigen combinations, we selected the top three antigens based on AUCs and assigned a score of 0 or 1 to each antigen result depending on whether it was below or above the cutoff level for the antigen; a positive responder was determined if any two antigen tests were positive with total score ≥ 2.

For logistic regression analysis, we selected 11 antigens as factors to construct the logistic regression model, including 64 LTBI individuals, and 62 active TB patients as the training set. We used a stepwise forward selection procedure to determine candidate biomarkers that contributed maximally to distinguish between LTBI and active TB. Stepwise procedures were guided by an F value probability of 0.05 for inclusion, and 0.10 for exclusion. The leave-one-out method was used to evaluate the model's generalization ability. Coefficients for antigens included in the final step were calculated. The accuracy of the established regression model was evaluated by ROC analysis. We selected independently 29 LTBI, and 30 active TB patients as the validation set.

Statistical analyses were performed using GraphPad Prism V5.0 (GraphPad Software, CA, USA) and SPSS version 17.0 (IBM, NY, USA).

RESULTS

Characteristics of the Study Population

In total, 319 subjects were enrolled in the final analysis, which consisted of 94 healthy controls, 113 LTBI subjects, and 112 active TB patients. Both demographic and clinical characteristics of all subjects in this study are shown in Table 1. In the healthy control, LTBI, and active TB groups, median ages were 40.5 years (range, 18-65 years), 45 years (range, 20-69 years), and 42 years (range, 18-65 years), respectively; the male to female ratios were

Table 1. Demographic Characteristics and Clinical Details of Study Subjects

Items	Health Controls	LTBI Subjects	Active TB Patients
Total number	94	113	112
Median age (range), years	40.5 (18-65)	45 (20-69)	42 (18-65)
Gender, male/female	44/50	46/67	53/59
Abnormal chest radiograph, n (%)	0 (0)	0 (0)	112 (100.0)
Bacteria positive, n (%)	nd	nd	112 (100.0)
TST results			
Induration < 5 mm, n (%)	94 (100.0)	0 (0)	nd
Induration 5-10 mm, n (%)	0 (0.0)	0 (0)	nd
Induration > 10 mm, n (%)	0 (0.0)	113 (100.0)	nd
T-SPOT results			
Positive, n (%)	0 (0)	113 (100.0)	89 (79.5)
Negative, n (%)	94 (100.0)	0 (0)	4 (3.6)
Unknown, n (%)	0 (0)	0 (0)	19 (16.9)
No immunosuppression, n (%)	94 (100.0)	113 (100.0)	112 (100.0)
HIV-negative, n (%)	94 (100.0)	113 (100.0)	112 (100.0)

Note. *n*, number of subjects; nd, not done.

44/50, 46/67, and 53/59, respectively, and there were no statistical differences among them with respect to age (P = 0.268) or gender (P = 0.600). For the T-spot TB assay, results for healthy controls and LTBI subjects were negative and positive, respectively. Furthermore, the T-spot TB assay detected 89 out of 93 patients with active TB, indicating a sensitivity of 95.7%. In addition, none of the subjects was infected with HIV or was treated with immunosuppressive medications.

Among these subjects, we randomly selected 20 each of LTBI and active TB patients as the screening population (screening using microarray), and 94 healthy controls, 93 LTBI individuals, and 92 active TB patients as the validation population (validation using indirect ELISA).

Screening of Serum Biomarkers

After normalizing and excluding the outlier microarrays, 18 and 19 microarray results were obtained from LTBI subjects and active TB patients, respectively. We further revealed that the concentrations of 152 Mtb antigen-specific IgG antibodies were significantly higher in active TB patients than in LTBI subjects based on a P value of < 0.05 and a fold-change (Log2) of > 0.18 (Supplementary Table S1, available in www.besjournal.com). Moreover, the differences in IgG reactivity against 152 antigens in each individual serum from LTBI individuals and active TB patients were subjected to hierarchical clustering, which distinguished between LTBI and active TB patients (Figure 1).

String Analysis

STRING analysis was performed to investigate the interactions of 152 candidate antigens (Figure 2). Among these antigens, 13 pairs, namely, Rv2499c *vs*. Rv2503c, Rv2502c *vs*. Rv1070c, Rv0400c *vs*. Rv0914c, Rv2908c *vs*. Rv2906c, Rv0400c *vs*. Rv1472c, Rv0632c *vs*. Rv0400c, Rv0914c *vs*. Rv2503c, Rv0400c *vs*. Rv1070c, Rv0914c *vs*. Rv1472c, Rv0914c *vs*. Rv0632c, Rv0914c *vs*. Rv1070c, Rv3501c *vs*. Rv0589c, and Rv2502c *vs*. Rv2499c, exhibited strong associations with the highest confidence (combined score > 0.9), implying that multiple Mtb antigens may collaborate to participate in resistance to host immune responses.

Classification of Candidate Antigens

The functional classification of 152 antigens was further analyzed using the TB database (Figure 3; Supplementary Table S2, available in www. besjournal.com). These Mtb antigens were involved in various biological processes. The top three of these processes with known functions were intermediary metabolism and respiration (27.63%), cell wall and cell processes (19.74%), and lipid metabolism (11.84%). In addition, large proportions (12.50%) of the antigens were conserved hypothetical proteins whose particular functions remain unknown.

Eleven Candidate Antigens Differing between Active TB and LTBI were Validated Using ELISA

To validate the potential capability of candidate antigens for discriminating LTBI from active TB, 11 antigens that had high fold-change values and could easily be expressed and purified were selected and further assessed by ELISA using 279 serum samples including 94 healthy controls, 93 LTBI patients, and 92 active TB patients. As expected, the 11 antigen-specific IgG levels were significantly higher in active TB patients than in LTBI individuals, which is consistent with microarray results (Figure 4).

Performance of 11 Antigens in Differentiating between Active TB and LTBI

The performance of 11 antigens in distinguishing between LTBI and active TB was evaluated using ROC curves, and their AUCs, specificities, and sensitivities were calculated (Table 2). Results showed that Rv2031c, Rv1408, and Rv2421c had higher AUC of 0.8520 (95% *CI:* 0.7954-0.9086), 0.8152 (95% *CI*: 0.7494-0.8809) and 0.7970 (95% *CI*: 0.7271-0.8668), respectively. Several other antigens, including Rv2716, Rv2002, Rv2097c, Rv0248c, Rv2026c, Rv0389, Rv2906c, and Rv2928, also distinguished between LTBI and active TB, and the results were statistically significant (P < 0.05).

Improved Diagnostic Performance by Antigen Combination

According to AUC values, we selected the top three antigens, namely Rv2031c, Rv1408, and Rv2421c to perform antigen combination analysis. Our results showed that the positive rate of antigen combination in the LTBI group (namely, specificity) was increased to 93.6%, which is higher than that obtained for any single antigen (Figure 5A). There was a statistically significant difference in the detection rate of Rv2421c *vs*. combination (P = 0.002), and Rv2031c *vs*. combination (P = 0.039). The positive rate of antigen combination in the active TB group (namely,
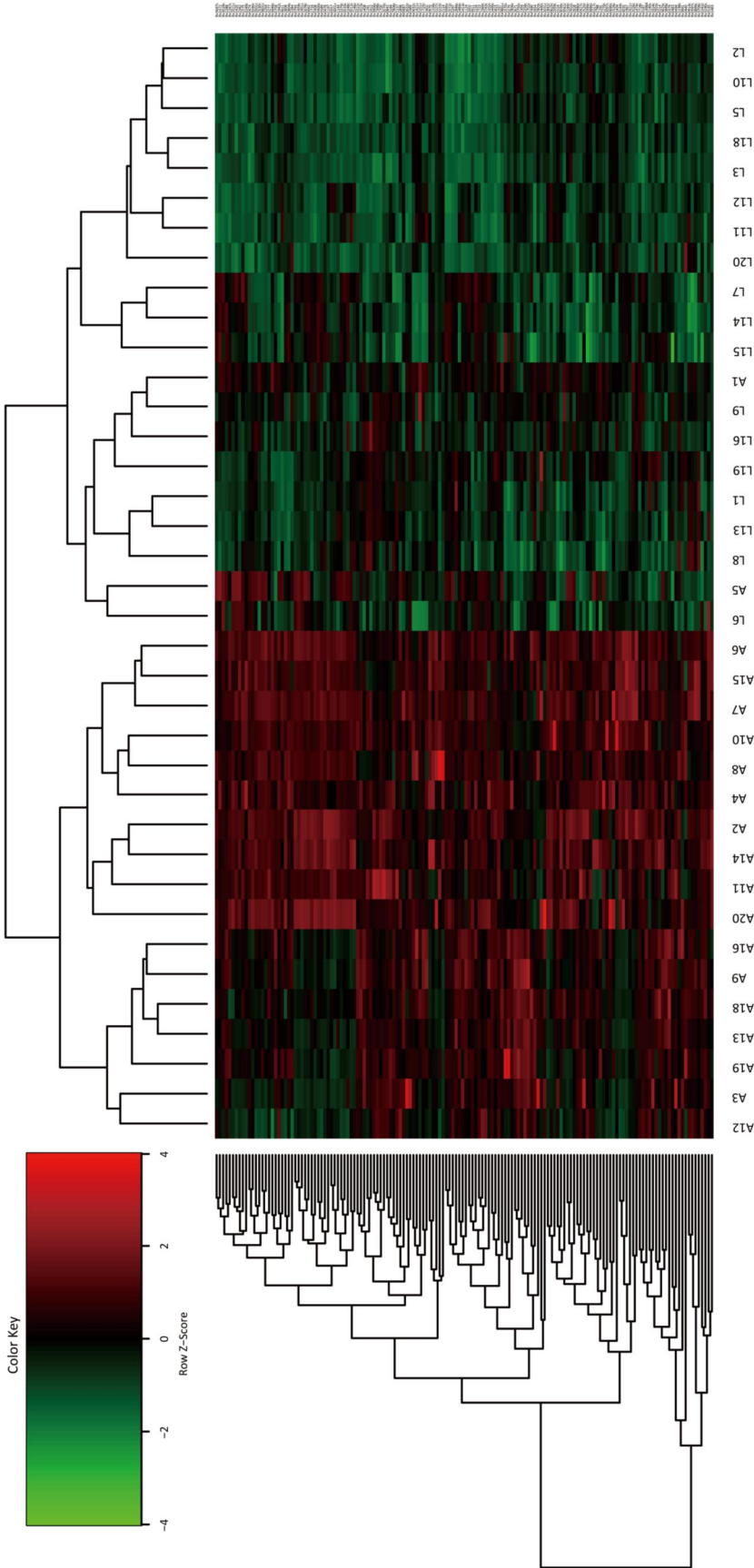
Figure 1. Hierarchical clustering of differentially expressed serum biomarkers. Each column indicates one serum sample and each row indicates one Mtb antigen. Data were normalized by Z score estimation, and the difference in serum reactivity was visualized as a color spectrum. A, active TB; L, latent TB infection.
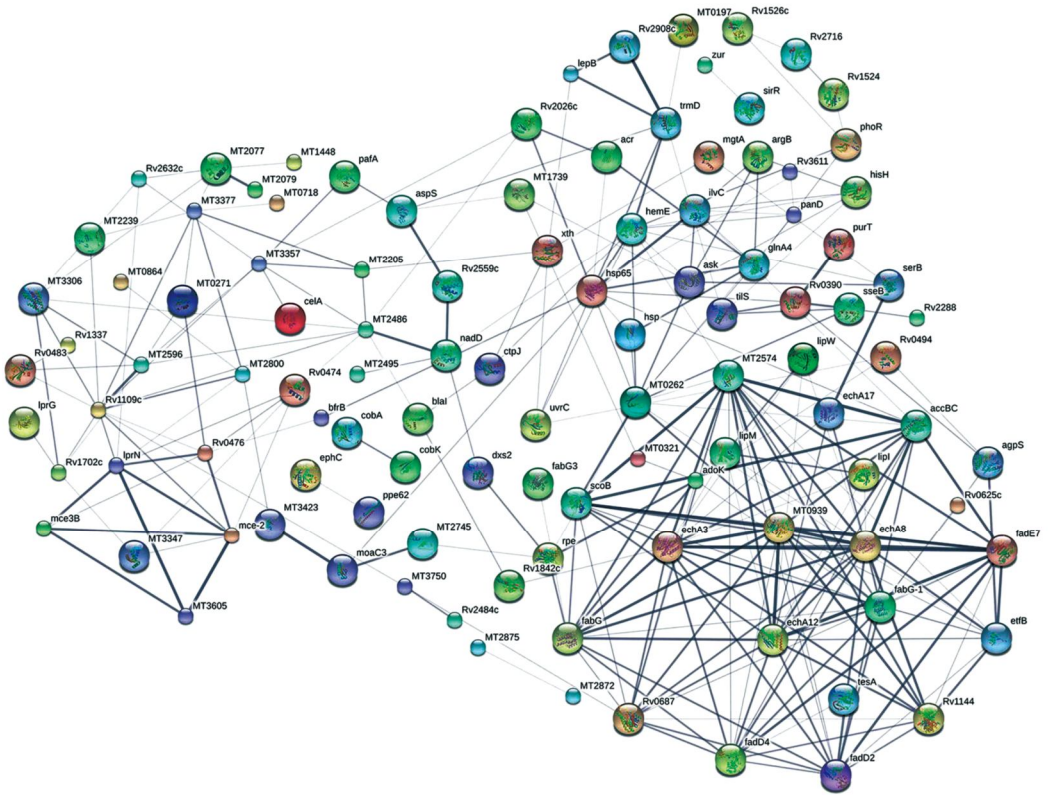
Figure 2. STRING analyses of candidate Mtb antigens. Interaction networks were visualized with a confidence level of 0.25. Stronger associations are represented by thicker lines.
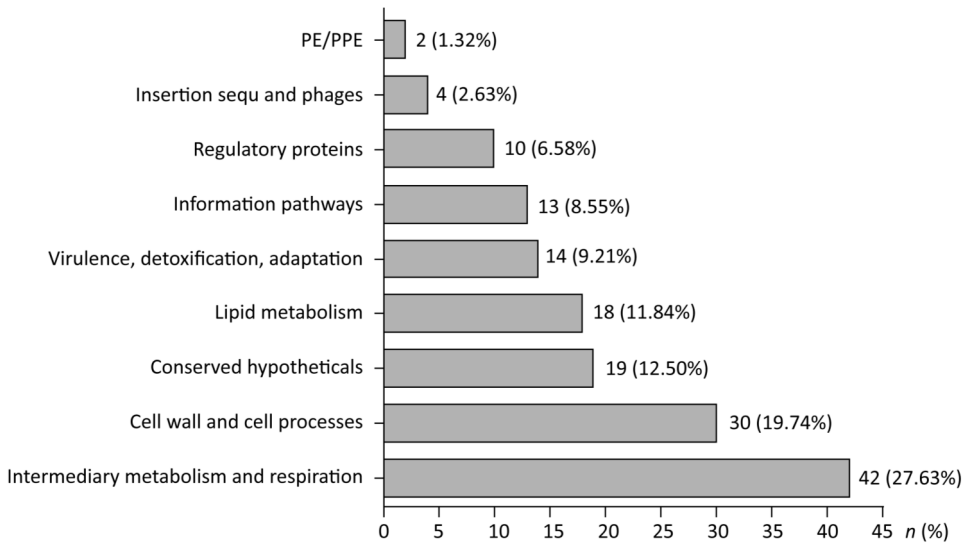

Figure 3. Functional category of candidate Mtb antigens. Analysis is based on the Sanger Institute database. N, number of antigens; PE/PPE, Pro-Glu/Pro-Pro-Glu.
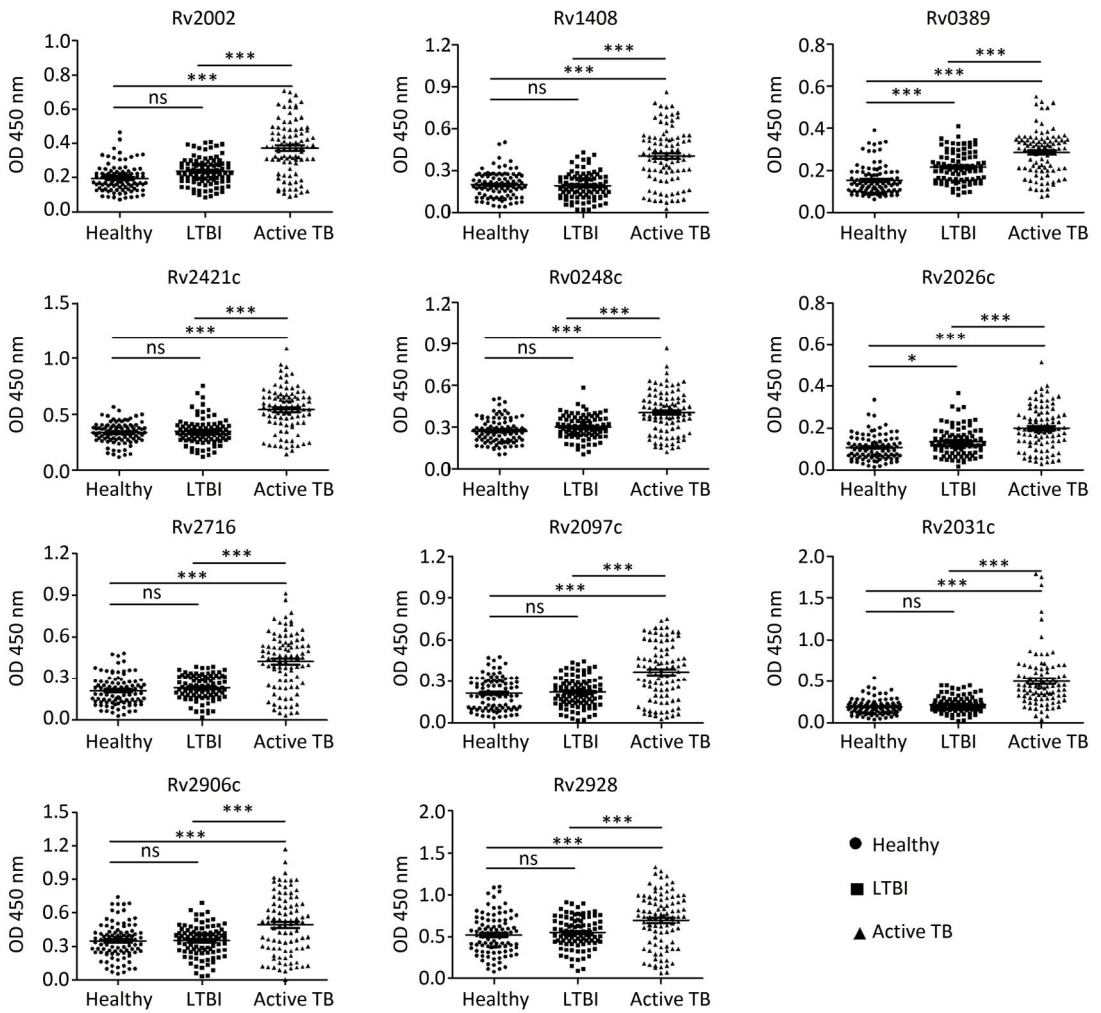
Figure 4. Scatter plots of IgG levels against 11 Mtb antigens in patients with active TB, LTBI, and healthy controls. Bars indicate means, and error bars indicate SEM. $^*P < 0.05$; $^{***}P < 0.001$; ns, not significant.

Table 2. AUCs for Discriminating between LTBI and Active TB Patients

Rv Number	Gene Name	AUC (95% *CI*)	*P* Value	Sensitivity (%, 95% *CI*)	Specificity (%, 95% *CI*)
Rv2031c	*hspX*	0.8520 (0.7954-0.9086)	< 0.0001	75.00 (64.89-83.45)	84.95 (76.03-91.52)
Rv1408	*rpe*	0.8152 (0.7494-0.8809)	< 0.0001	75.00 (64.89-83.45)	87.10 (78.55-93.15)
Rv2421c	*nadD*	0.7970 (0.7270-0.8668)	< 0.0001	77.17 (67.25-85.28)	82.80 (73.57-89.83)
Rv2716	*Rv2716*	0.7931 (0.7212-0.8650)	< 0.0001	66.30 (55.70-75.83)	97.85 (92.45-99.74)
Rv2002	*fabG3*	0.7765 (0.7047-0.8484)	< 0.0001	73.91 (63.71-82.52)	83.87 (74.80-90.68)
Rv2097c	*Rv2097c*	0.7184 (0.6415-0.7954)	< 0.0001	60.87 (50.14-70.88)	80.65 (71.15-88.11)
Rv0248c	*Rv0248c*	0.7176 (0.6395-0.7956)	< 0.0001	69.57 (59.10-78.73)	73.12 (62.92-81.79)
Rv2026c	*Rv2026c*	0.7010 (0.6237-0.7782)	< 0.0001	67.39 (56.82-76.80)	69.89 (59.50-78.97)
Rv0389	*purT*	0.7021 (0.6254-0.7789)	< 0.0001	58.70 (47.95-68.87)	82.80 (73.57-89.83)
Rv2906c	*trmD*	0.6605 (0.5781-0.7430)	0.0002	52.17 (41.50-62.70)	89.25 (81.11-94.72)
Rv2928	*tesA*	0.6584 (0.5760-0.7409)	0.0002	42.39 (32.15-53.14)	92.47 (85.11-96.92)

Note. 95% *CI*: 95% confidence interval.

sensitivity) was increased to 81.5%, which was slightly higher than that observed for each antigen, although it was not statistically significant (Figure 5B). In addition, the positive rate of each antigen did not show any statistical difference with respect to the LTBI or the active TB group.

Logistic Regression Analysis of Candidate Antigens for Distinguishing between LTBI and Active TB

Logistic regression analysis was used to enhance diagnostic ability. Our results showed that Rv1408, R0248, Rv2026c, Rv2716, Rv2031c, Rv2928, and Rv2121c were the major predictive antigens for differentiating between LTBI and active TB patients. Therefore, a regression model Logit $P = -6.28 + 16.92 \times Rv1408 - 11.26 \times R0248 + 19.82 \times Rv2026c + 25.49 \times Rv2716 + 11.35 \times Rv2031c - 23.07 \times Rv2928 + 11.55 \times Rv2121c$ was established. In the training set, the AUC of the model was 0.9844 (95% CI: 96.62%-1.003%), with sensitivity and specificity of 96.77% (95% CI: 88.83%-99.61%) and 93.75% (95% CI:

84.76%-98.27%), respectively (Figure 6A). In the validation set, the AUC of the model was 0.9632 (95% CI: 91.40%-1.012%), with sensitivity and specificity of 93.33% (95% CI: 77.93%-99.18%) and 93.1% (95% CI: 77.23%-99.15%), respectively (Figure 6B).

DISCUSSION

Since Mtb is an intracellular pathogen, cell-mediated immunity has always been considered as the predominant defense mechanism against Mtb infection. However, accumulating evidence suggests that B cell-mediated humoral immune response may play an important role in Mtb infection and is related to the progression of latent infection to active TB disease[15-18]. Previous studies showed several Mtb antigens such as ESAT-6 (Rv3875), CFP-10 (Rv3874), Ag85B (Rv1886c), Hsp16.3 (16-kD, Rv2031c), 38 kD (Rv0934), LAM, Rv2029c, Rv2628, and Rv1813c have previously been reported to be able to discriminate
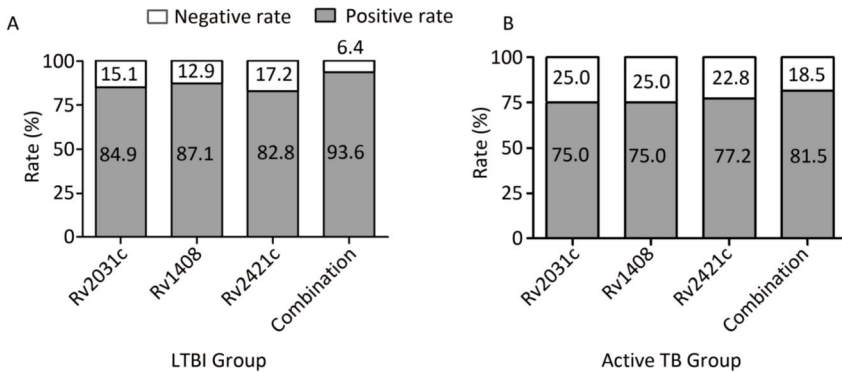

Figure 5. Distribution of positive and negative responders to Rv2031c, Rv1408, Rv2421c antigens, and a combination of these antigens. (A) Latent TB infection group. (B) Active tuberculosis group. For antigen combinations, a positive responder was determined if any two antigens yielded a positive result. Detection rates were compared using McNemar's test.
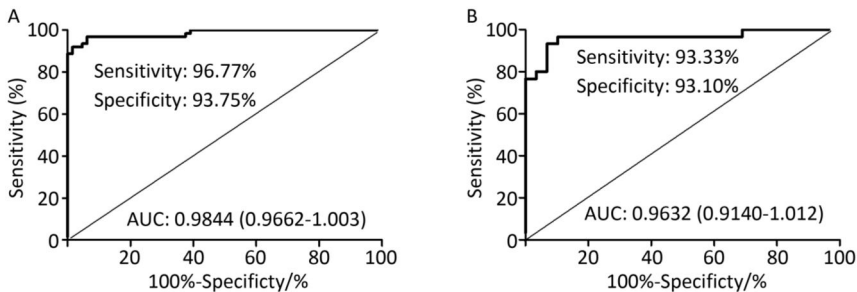

Figure 6. ROC curve analysis of logistic regression model. (A) Training set. (B) Validation set. AUC, sensitivity, and specificity are shown.

between LTBI and active TB[19-21]. Recently, Lu et al.[22] demonstrated that antibody glycosylation patterns could also distinguish between LTBI and active TB. In our study, we aimed to screen novel serum biomarkers for discrimination between LTBI and active TB at the systemic level using an Mtb proteome microarray containing 4,262 antigens. We found that the concentrations of 152 Mtb antigen-specific IgG antibodies were higher in the active TB group than in the LTBI group. This result provided potential serum biomarkers and data that may facilitate a better understanding of differences in the host humoral immune response between LTBI and active TB. Among previously reported and commonly used antigens, only Rv2031c was included in our results. This may be because the protein antigens used in previous studies were all produced using an *Escherichia coli*-based cell-free transcription system and without further purification. However, in our study, the antigens spotted on the Mtb proteome microarray were all produced using a eukaryotic yeast expression system. The use of a eukaryotic system results in the expressed protein antigens having many more posttranslational modifications than those produced using a prokaryotic system. In the present study, the antigens were obtained by gently using affinity purification to maintain the activity that is necessary for functional studies[16]. This is important since specific features of these antigens may have a great influence on the antibody response. In addition, the enrollment of subjects with a different infection time or status and with or without anti-TB treatment may impact the results. These reasons may underlie the inconsistency of current approaches for the serological diagnosis of TB.

In our validation experiments using the 11 candidate antigens and ELISA, we used healthy controls to ensure the accuracy of the test, i.e., the levels of antigen-specific IgG antibodies must be higher in both the LTBI and active TB groups than in the healthy control group. It is noteworthy that Rv0389, Rv2002, and Rv2026c showed significant differences among the healthy control, LTBI, and active TB groups, suggesting that these antigens may elicit stronger humoral immune responses despite low bacterial load in LTBI individuals. Therefore, Rv0389, Rv2002, and Rv2026c antigens could be potential biomarkers for diagnosing Mtb infection and determining disease status. However, further studies in this regard are warranted.

Antigens assessed in this study (Rv2031c, Rv1408, and Rv2421c) could not discriminate between LTBI subjects and healthy controls. However, they could distinguish between LTBI and active TB with sufficient accuracy, indicating that these antigens may have potential value by offering improved predictive accuracy when used in combination with IGRAs. Rv2031c is involved in virulence, detoxification, and adaptation to biological processes and is predominantly expressed by Mtb during latency[23,24]. In contrast, both Rv1408 and Rv2421c are involved in intermediary metabolism and respiration, which are required for the long-term survival of Mtb. Specifically, Rv1408 is ribulose-phosphate 3-epimerase (rpe), which catalyzes the reversible epimerization of D-ribulose 5-phosphate to D-xylulose 5-phosphate. Previous studies showed that Rv1408 expression was downregulated during nitrogen starvation, suggesting that it was involved in nitrate/nitrite metabolism[25]. Rv2421c is nicotinate-nucleotide adenylyltransferase, which transfers phosphate groups during NAD biosynthesis and salvage. Cloete et al. reported that Rv2421c was upregulated during starvation or dormancy by hypoxia[26]. This is the first study to demonstrate that Rv1408 and Rv2421c antigens might be used for distinguishing LTBI from active TB patients.

Since significant heterogeneity exists among individuals and various regions, a single antigen can never achieve satisfactory performance for the diagnosis of TB. Deng et al.[14] have demonstrated that the combination of several serum biomarkers showed better performance in distinguishing between active TB patients and recovered individuals, with 78.4% diagnostic accuracy, 80.2% sensitivity, and 79.0% specificity. He et al.[27] reported that the combination of Rv3618 with the 38 kD 38E6 could improve the sensitivity of TB diagnosis. In this study, we combined several antigens to improve diagnostic ability. We selected the top three antigens with high AUC values to conduct antigen combination analysis. Our results showed that the specificity of the antigen combination was significantly higher than those of single antigens. We also used logistic regression analysis to develop a regression model containing seven antigens and cross-validated the results using another independent sample. Once again, our results showed that both specificity and sensitivity were significantly better than those observed with any single antigen. Although logistic regression analysis is rarely used in studies on TB, it has been applied in several other studies[28,29]. However, it is noteworthy that the

seven antigens included in the regression model were not the top seven out of the 11 antigens according to AUC values. Some of them had relatively lower AUCs but higher specificity than other antigens, which when combined with other antigens of higher sensitivity, may result in a mutual complementation pattern for distinguishing LTBI from active TB.

However, this study still has certain limitations. The primary limitation is that the sputum smear or bacterial culture of active TB patients were Mtb-positive and bacteria-negative TB patients were not considered. In addition, the sample size in the logistic regression model construction and validation phases was small. Therefore, further studies for verifying the diagnostic performance of the regression model with larger sample size, including bacteria-negative TB patients, are required in the future.

In conclusion, we have revealed that a variety of Mtb antigens could elicit stronger serum IgG responses in active TB patients when compared with LTBI subjects using proteome microarrays, which may be useful for understanding host humoral immune responses in individuals with different Mtb infection statuses. Furthermore, several antigens were further validated. Each of these antigens exhibited a moderate diagnostic performance in discriminating between active TB and LTBI while using the top three antigens in combination or using a logistic regression model established with seven antigens could improve the diagnostic sensitivity and specificity. Of these two approaches, the logistic regression model with seven antigens had the best performance, suggesting that it may serve as an index for distinguishing between active TB and LTBI.

REFERENCES

1. World Health Organization. Global tuberculosis report 2016. http://www.who.int/tb/publications/global_report/en. [2017-02-01].

2. Skeiky YA, Sadoff JC. Advances in tuberculosis vaccine strategies. Nat Rev Microbiol, 2006; 4, 469-76.

3. World Health Organization. Guidelines on the management of latent tuberculosis infection. 2015. http://www.who.int/entity/tb/publications/latent-tuberculosis-infection/en/. [2017-2-1].

4. Mack U, Migliori GB, Sester M, et al. TBNET. LTBI: latent tuberculosis infection or lasting immune responses to M. tuberculosis? A TBNET consensus statement. Eur Respir J, 2009; 33, 956-73.

5. Gupta A, Kaul A, Tsolaki AG, et al. Mycobacterium tuberculosis: Immune evasion, latency and reactivation. Immunobiology, 2012; 217, 363-74.

6. Goosen WJ, Cooper D, Miller MA, et al. IP-10 Is a Sensitive Biomarker of Antigen Recognition in Whole Blood Stimulation Assays Used for the Diagnosis of Mycobacterium bovis Infection in African Buffaloes (Syncerus caffer). Clin Vaccine Immunol, 2015; 22, 974-8.

7. Harboe M. Antigens of PPD, old tuberculin, and autoclaved Mycobacterium bovis BCG studied by crossed immunoelectrophoresis. Am Rev Respir Dis, 1981; 124, 80-7.

8. Liebeschuetz S, Bamber S, Ewer K, et al. Diagnosis of tuberculosis in South African children with a T-cell-based assay: a prospective cohort study. Lancet, 2004; 364, 2196-203.

9. Bianchi L, Galli L, Moriondo M, et al. Interferon-gamma release assay improves the diagnosis of tuberculosis in children. Pediatr Infect Dis J, 2009; 28, 510-4.

10. Andersen P, Munk ME, Pollock JM, et al. Specific immune-based diagnosis of tuberculosis. Lancet, 2000; 356, 1099-104.

11. Suzukawa M, Akashi S, Nagai H, et al. Combined Analysis of IFN-γ, IL-2, IL-5, IL-10, IL-1RA and MCP-1 in QFT Supernatant Is Useful for Distinguishing Active Tuberculosis from Latent Infection. PLoS One, 2016; 11, e152483.

12. Anderson BL, Welch RJ, Litwin CM. Assessment of three commercially available serologic assays for detection of antibodies to Mycobacterium tuberculosis and identification of active tuberculosis. Clin Vaccine Immunol, 2008; 15, 1644-9.

13. Sardella IG, Singh M, Kumpfer S, et al. Evaluation of Lionex TB kits and mycobacterial antigens for IgG and IgA detection in cerebrospinal fluid from tuberculosis meningitis patients. Mem Inst Oswaldo Cruz, 2010; 105, 722-8.

14. Deng J, Bi L, Zhou L, et al. Mycobacterium Tuberculosis Proteome Microarray for Global Studies of Protein Function and Immunogenicity. Cell Rep, 2014; 9, 2317-29.

15. Chan J, Mehta S, Bharrhan S, et al. The role of B cells and humoral immunity in Mycobacterium tuberculosis infection. Semin Immunol, 2014; 26, 588-600.

16. Chen T, Blanc C, Eder AZ, et al. Association of Human Antibodies to Arabinomannan With Enhanced Mycobacterial Opsonophagocytosis and Intracellular Growth Reduction. J Infect Dis, 2016; 214, 300-10.

17. Phuah J, Wong EA, Gideon HP, et al. Effects of B Cell Depletion on Early Mycobacterium tuberculosis Infection in Cynomolgus Macaques. Infect Immun, 2016; 84, 1301-11.

18. Torrado E, Fountain JJ, Robinson RT, et al. Differential and site specific impact of B cells in the protective immune response to Mycobacterium tuberculosis in the mouse. PLoS One, 2013; 8, e61681.

19. Zhang C, Song X, Zhao Y, et al. Mycobacterium tuberculosis Secreted Proteins As Potential Biomarkers for the Diagnosis of Active Tuberculosis and Latent Tuberculosis Infection. J Clin Lab Anal, 2015; 29, 375-82.

20.Hur YG, Kim A, Kang YA, et al. Evaluation of antigen-specific immunoglobulin g responses in pulmonary tuberculosis patients and contacts. J Clin Microbiol, 2015; 53, 904-9.

21.Bai XJ, Liang Y, Yang YR, et al. Potential novel markers to discriminate between active and latent tuberculosis infection in Chinese individuals. Comp Immunol Microbiol Infect Dis, 2016; 44, 8-13.

22.Lu LL, Chung AW, Rosebrock TR, et al. A Functional Role for Antibodies in Tuberculosis. Cell, 2016; 167, 433-43.

23.Hu Y, Movahedzadeh F, Stoker NG, et al. Deletion of the Mycobacterium tuberculosis alpha-crystallin-like hspX gene causes increased bacterial growth in vivo. Infect Immun, 2006; 74, 861-8.

24.Mushtaq K, Sheikh JA, Amir M, et al. Rv2031c of Mycobacterium tuberculosis: a master regulator of Rv2028-Rv2031 (HspX) operon. Front Microbiol, 2015; 6, 351.

25.Williams KJ, Jenkins VA, Barton GR, et al. Deciphering the metabolic response of Mycobacterium tuberculosis to nitrogen stress. Mol Microbiol, 2015; 97, 1142-57.

26.Cloete R, Oppon E, Murungi E, et al. Resistance related metabolic pathways for drug target identification in Mycobacterium tuberculosis. Bmc Bioinformatics, 2016; 17, 75.

27.He XY, Li J, Hao J, et al. Assessment of five antigens from Mycobacterium tuberculosis for serodiagnosis of tuberculosis. Clin Vaccine Immunol, 2011; 18, 565-70.

28.Zhu C, Yang H, Geng Q, et al. Association of oxidative stress biomarkers with gestational diabetes mellitus in pregnant women: a case-control study. PLoS One, 2015; 10, e0126490.

29.Dessau RB, Ejlertsen T, Hilden J, et al. Simultaneous use of serum IgG and IgM for risk scoring of suspected early Lyme borreliosis: graphical and bivariate analyses. APMIS, 2010; 118, 313-23.