Screening and Assessing 11 *Mycobacterium tuberculosis* Proteins as Potential Serodiagnostical Markers for Discriminating TB Patients from BCG Vaccinees

Guoqiang Zhang[1,2#], Lingxia Zhang[1,3#], Mingcheng Zhang[1], Linlin Pan[1], Fengyu Wang[1,2], Jun Huang[1], Guoli Li[3], Jun Yu[1], and Songnian Hu[1*]

[1] *CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China;* [2] *Graduate University of Chinese Academy of Sciences, Beijing 100049, China;* [3] *The Second Hospital Affiliated to the General Hospital of the People's Liberation Army, Beijing 100091, China.*

[#]Equal contribution.

*Corresponding author. E-mail: husn@big.ac.cn

DOI: 10.1016/S1672-0229(08)60039-X

Purified protein derivative (PPD) skin tests often yield poor specificity, so that to develop new serological antigens for distinguishing between *Mycobacterium tuberculosis* infection and Bacille Calmette-Guerin (BCG) vaccination is a priority, especially for developing countries like China. We predicted the antigenicity for selected open reading frames (ORFs) based on the genome sequences of *M. tuberculosis* H37Rv and *M. bovis* BCG, as well as their functions and differences of expression under different stimulus. The candidate ORFs were cloned from H37Rv sequences and expressed as recombinant proteins in *Escherichia coli*. We studied the serodiagnostic potential of 11 purified recombinants by using enzyme-linked immunosorbent assay (ELISA) and involving a cohort composed of 58 TB patients (34 males and 24 females), 8 healthy volunteers and 50 PPD-negative individuals before and after BCG vaccination. For all the 11 antigens, the median OD values for the sera from TB patients were statistically significantly higher than those for PPD-negative individuals before or after BCG vaccination ($P<0.01$). They had at least 92% specificity in healthy controls and six seroantigens (Rv0251c, Rv1973, Rv2376c, Rv2537c, Rv2785c and Rv3873A) were never reported with seroantigenicities previously. Thus the approach combining comparative genomics, bioinformatics and ELISA techniques can be employed to identify new seroantigens distinguishing *M. tuberculosis* infection from BCG vaccination.

Key words: comparative genomics, *M. tuberculosis*, BCG, seroantigen, ELISA

Introduction

Tuberculosis (TB) is an infectious bacterial disease caused by *Mycobacterium tuberculosis*. Based on surveillance and survey data, the World Health Organization (WHO) estimated that 9.27 million new cases of TB occurred in 2007 globally (139 per 100,000 population). An estimated 1.32 million HIV-negative people died from TB in 2007, and there were an additional 456,000 TB deaths among HIV-positive people (*1*).

Currently, TB diagnosis largely depends on clinical and radiographic (X-ray) examination validated by sputum smear microscopy and bacterial culture. Culture is not practical as a first-line defense due to long cultivation period. Present confirmation of diagnosis relies on sputum smear examination, achieving a sensitivity ranging from 10% to 20%, and low specificity. To make the diagnosis of TB more accurate, fast and convenient, new diagnostic techniques, including *M. tuberculosis* detection in clinical specimens by PCR and immune reactions based on cell-mediated-immune (CMI) or humoral immune response, have been investigated. A serological test, unlike the CMI-based diagnostic assays requiring living cells, is an attractive diagnostic method for its con-

venience, robustness and easy implementation in developing countries. However, no commercially available serological tool has so far shown sufficient sensitivity and specificity to early detect *M. tuberculosis* infection and differentiate infection from Bacille Calmette-Guerin (BCG) vaccination (*2*). The reason is generally believed to be the great heterogeneous antibody response in TB patients (*3*). A necessary improvement in sensitivity can be achieved by combining the best antigens, through the use of both cocktails of single proteins and genetically engineered fusion molecules containing several antigens (*4*, *5*).

In recent years, MTB48 (*6*), ESAT-6, CFP-10 (*7*, *8*) and other antigens were discovered by the techniques in screening expression library of *M. tuberculosis* and by comparative genomics and proteomics (*9–12*). Based on these observations, we hypothesized that candidate antigens, with minimum similarity to BCG or presumably drastically regulated under different stimulus, may provide more specific targets to measure humoral immunity. We expressed 17 candidate antigens mainly based on *M. tuberculosis* H37Rv sequence in *Escherichia coli* and serologically evaluated 11 of them among TB patients and purified protein derivative (PPD)-negative individuals before and after BCG vaccination. Between the sera from TB patients and PPD-negative individuals before or after BCG vaccination, the levels of antibodies against all the 11 assayed antigens were statistically different ($P<0.01$). All the 11 assayed antigens had at least 92% specificity in healthy controls and six seroantigens (Rv0251c, Rv1973, Rv2376c, Rv2537c, Rv2785c and Rv3873A) were newly found.

Results

Selection of high antigenic proteins on the basis of comparing proteins between H37Rv and BCG_P

The tBLASTn result showed that the identities of 174 genes between H37Rv and BCG_P genomes were less than 60%. Among them, 49 genes have signal peptide or transmembrane structure and 55 are annotated to MCE family or PE/PPE family, which is supposed to play important roles in cell invasion or immune responses (*13–17*). Therefore, we first selected to clone and express 12 candidates belonging to these two kinds of genes with high antigenic index over 3.0 predicted by BEPITOPE software (**Table 1**). Additionally, we also picked out 4 reportedly promising antigens, 2 heat shock proteins (HSPs) and 2 enzymes associated with metabolism to clone and express (Table 1). Based on the characteristic structure and the predicted epitopes of the PPE proteins, we dissected Rv3873 into the conserved N-terminal (PPE region: Rv3873A) and the diverse C-terminal (MPTR region: Rv3873B). In addition, we dissected Rv1563c into Rv1563cA and Rv1563cB for the same reason.

Selection of H37Rv proteins recognized serologically by sera of TB patients

In the total 22 proteins to be purified, 3 genes (Rv1968, Rv1243c and Rv3590c) failed to be cloned and 2 others (Rv1987 and Rv2875) failed to be purified. Finally, we successfully cloned and purified 17 H37Rv proteins (**Table 2**). The serological antigenicities of most of them have not been previously identified by proteomic or genomic approaches. We also included the PPD protein, because of its homologs among several other mycobacteria, to serologically evaluate its sensitivity and specificity among the healthy, BCG-vaccinated and *M. tuberculosis*-infected individuals.

During the first round of screening step, we attempted to identify a broad panel of targets by using sera from 8 selected TB patients (including one with kidney TB and another with liver TB) who all had high titers of antibody against PPD from *M. tuberculosis*. The levels of specific antibodies in these TB blood samples were compared with those in 8 healthy controls. Generally, the antibody responses to the antigens could be grouped into three different categories. The antibody levels in the sera of TB patients were higher than those in healthy controls for the proteins in category 1, contrast to those in category 2, in which the antibody levels in the sera from the patients were lower than those in healthy controls. The proteins in category 3 were characterized by low antibody responses to the sera from patients as well as to healthy controls. Only antigens in category 1 with marked difference (a mean OD difference of more than 0.2) between the signal in patients and controls were selected for further analysis (Table 2). Eleven antigens (Rv0251c, Rv1973, Rv1984c, Rv2031c, Rv2185c, Rv2376c, Rv2537c, Rv2785c, Rv2878c, Rv0934 and Rv3873A) satisfying this criterion were selected. Among them, there is no report for the serological antigenicities of Rv0251c, Rv1973, Rv2376c, Rv2537c, Rv2785c and Rv3873A proteins.

Table 1 Assumed characteristics and identities of candidate antigens in H37Rv

Rv No. of gene	Gene symbol*[1]	Identity (%) with BCG_P (tBLASTn)*[2]	Signal peptide*[3]	Transmembrane*[3]	Methods for detecting protein
Rv0251c	hsp	100	−	−	*in vivo* (*33*)
Rv0934	pstS1	99.47	+	−	Culture *in vitro* (*34*)
Rv1243c*[4]	PE_PGRS23	52.39	+	−	Computational prediction
Rv1563c*[6]	treY1	98.73	−	−	*in vivo* (*35*)
Rv1966	mce3A	36	+	+	*in vivo* (*36*)
Rv1968*[4]	mce3C	44	+	+	Computational prediction
Rv1971	mce3F	39	+	+	Computational prediction
Rv1973	possible mce	49	+	+	Computational prediction
Rv1984c	cfp21	49.78	+	−	Culture *in vitro* (*37*)
Rv1987*[5]	possible chitinase	37.74	+	+	Computational prediction
Rv2031c	hspX	99.31	−	−	*in vivo* (*20*)
Rv2185c	TB16.3	99.31	−	−	Culture *in vitro* (*38*)
Rv2376c	cfp2	29.63	+	−	Culture *in vitro* (*39*)
Rv2537c	aroD	34.33	−	+	Computational prediction
Rv2785c	rpsO	34.33	−	−	Computational prediction
Rv2875*[5]	mpt70	70.27	+	−	Culture *in vitro* (*40*)
Rv2878c	mpt53	100	+	+	Culture *in vitro* (*41*)
Rv3590c*[4]	PE_PGRS58	52	+	−	Computational prediction
Rv3804c	fbpA	100	+	+	Culture *in vitro* (*42*)
Rv3873*[6]	PPE68	45.65	+	−	*in vivo* (*15*)

*[1]Annotations for gene symbols are shown according to the database in NCBI.
*[2]All BLAST reports were performed within local server, and all the reference sequences of proteins in H37Rv and the sequence of BCG_P genome were downloaded from NCBI; tBLASTn results were reported in September 2005.
*[3]The "+" and "−" marks in "Signal peptide" column represent the protein with and without signal peptide; the "+" and "−" marks in "Transmembrane" column represent the protein with and without transmembrane structure.
*[4]The genes were not successfully cloned.
*[5]The recombinant proteins were not successfully purified.
*[6]The genes were dissected into two fragments to clone and express.

Establishment of the cutoffs and evaluation of the 11 serological antigens with sera from TB patients and BCG-vaccinated individuals

We evaluated 11 antigens and PPD with a panel of blood samples collected from 50 TB patients (including 6 extra pulmonary TB patients) and 50 PPD-negative individuals before and after BCG vaccination. For all the 11 antigens and PPD, the median OD values for the sera from TB patients were statistically significantly higher than those for PPD-negative individuals before or after BCG vaccination ($P<0.01$). For Rv1973, Rv2785c and PPD, the median OD values for the sera from PPD-negative individuals after BCG vaccination were statistically higher than those before BCG vaccination while, for Rv2185c, Rv2537c and Rv0934, the median OD values for the sera from PPD-negative individuals after BCG vaccination were statistically lower than those before BCG vaccination ($P<0.05$) (data not shown). The cutoffs for each antigen were established, and the sensitivity and specificity for each individual antigen were calculated (**Table 3**). The most frequently recognized antigens were Rv1973 and Rv2031c, which were recognized by 38% of the sera from the TB patients tested, while Rv1973 has a recognition frequency of 16% for BCG-vaccinated population. All antigens had a specificity of at least 92% for healthy controls.

Discussion

During recent years, antigens with performances similar to or better than that of the Rv0934 (38 kDa) protein have been identified (*5*). To identify additional potentially promising serological antigens, we combined comparative genomics, bioinformatics and enzyme-linked immunosorbent assay (ELISA) techniques to identify new seroantigens. Although we

Table 2 Serological screening of 17 recombinant *M. tuberculosis* proteins

Protein	Mean OD_{TB}	Mean OD_{Health}	ΔOD value*[1]	Selection*[2]	Reference
Rv0251c	0.963	0.519	0.444	+	NR*[3]
Rv0934	1.006	0.774	0.232	+	(*5*)
Rv1563cA	0.172	0.129	0.043	−	NR
Rv1563cB	0.274	0.309	−0.035	−	NR
Rv1966	0.220	0.134	0.086	−	NR
Rv1971	0.336	0.243	0.093	−	NR
Rv1973	0.419	0.184	0.235	+	NR
Rv1984c	0.493	0.275	0.218	+	(*5*)
Rv2031c	1.202	0.597	0.604	+	(*43*)
Rv2185c	1.119	0.644	0.475	+	(*5*)
Rv2376c	0.665	0.318	0.347	+	NR
Rv2537c	0.645	0.417	0.228	+	NR
Rv2785c	0.639	0.378	0.261	+	NR
Rv2878c	0.636	0.399	0.237	+	(*44*)
Rv3804c	0.164	0.299	−0.135	−	(*45*)
Rv3873A	0.319	0.094	0.225	+	NR
Rv3873B	0.215	0.163	0.052	−	NR
PPD	0.974	0.204	0.770	+	(*2*)

*[1]The change in OD (ΔOD) value is calculated as the mean OD for the healthy control subtracted from the mean OD for the TB patient. *[2]All antigens with ΔOD value >0.2 were selected ("+"). *[3]NR, not reported with seroantigenicity previously.

Table 3 Sensitivities and specificities of 11 recombinant proteins and PPD with sera from 50 TB patients and 50 PPD-negative individuals before and after BCG-vaccination

Protein	Cutoff*[1]	Sensitivity (%)			Specificity (%)*[2]
		TB patients	BCG	Controls	
Rv0251c	0.239	34	4	2	98
Rv0934	0.185	26	4	2	98
Rv1973	0.445	38	16	8	92
Rv1984c	0.570	32	4	4	96
Rv2031c	0.071	38	6	0	100
Rv2185c	0.826	18	0	2	98
Rv2376c	0.563	24	4	6	94
Rv2537c	0.147	28	0	0	100
Rv2785c	0.270	16	2	0	100
Rv2878c	0.511	16	4	2	96
Rv3873A	0.492	24	6	2	98
Rv0934+Rv0251c	0.656	70	4	4	96
PPD	0.334	60	20	2	98

*[1]Cutoff = Mean OD_{Health} + 2.576×SD. *[2]Specificity = 100% − sensitivity (for the controls).

purified 17 recombinant proteins, we dropped out 6 proteins after the first round of evaluation and only further evaluated other 11 with the sera from TB patients, healthy controls and BCG-vaccinated populations.

Six antigens (Rv0251c, Rv1973, Rv2376c, Rv2537c, Rv2785c and Rv3873A) were novel in the 11 evaluated antigens. Among the 6 novel antigens, Rv3873A and Rv1973 are worthy of noting. The 24% sensitivity of Rv3873A showed that the PPE gene

family can elicit strong humoral immune responses, although it is generally believed that PE/PPE fragment elicits T-cell immune responses in host organisms (*14–19*). It is also interesting that the sensitivity of Rv1973 antigen in BCG-vaccinated populations reaches 16% (the highest among the 11 evaluated antigens) given that its identity between H37Rv and BCG_P is only 49.47%. We attributed its 16% sensitivity to the polypeptide (ALLSYRPDTVQHDLESARSRLTG) in the high antigenic index region, sharing 60.9% identity with that (ALLSYSPDTLDQDFATARSHLAG) in BCG_P genome. Furthermore, from the great difference of the sensitivities of Rv1973 and PPD between the healthy controls and BCG-vaccinated individuals, we were tempted to imagine that the new serological antigens may be discovered to establish methods serologically monitoring the status of BCG vaccination timely (*20–22*), which was supposed to be extended to 10 to 15 years (*23*).

Six proteins have high predicted antigenicities but showed no significant differences in sera between TB patients and healthy controls. One reason may be that these proteins, like the proven Rv3804c (antigen 85A) (*24*, *25*), mainly elicit cellular immunity instead of humoral immunity. This is also the reason why these proteins are worthy of assessing their abilities to elicit cellular immunity. Another reason may be the highly specific humoral immunological response caused by its diversity (*18*, *19*), like PE/PPE protein family. The C-terminals of PE/PPE proteins reportedly mount on humoral immune response and their sequences are highly diverse. During the infection within the bodies of different individuals, the antibodies may be produced to be mainly against the different member of PE/PPE protein family. This speculation is substantiated, in our later experiment, with the 1.02 and 0.915 OD for the two sera recognized by Rv3873B protein and 0.434 OD as its cutoff, while its mean OD difference between TB patients and healthy controls is only 0.052.

In our study, all the 11 antigens can statistically differentiate TB patients from the BCG-vaccinated in terms of the median OD values between the antibody levels in the sampled sera. However, compared with other studies, the 16% to 38% sensitivities of the proteins in our study are lower. This may be resulted from the following reasons. First, the same antigen of different sources has different sensitivities in various studies (*23*, *26*, *27*). For example, the sensitivity of Rv0934, a well described lipoglycoprotein, ranges from 16% to 94% in different papers. Furthermore, the different post-translational modification between recombinant proteins and native antigens extracted from a *M. tuberculosis* complex may affect the humoral immune response. Additionally, the various living regions of the TB patients can also contribute to the different sensitivity of the same antigen. It is well known that China is an unevenly developing country with a large area, complex genetic background and over 90% immunization rate of infants with BCG from 1990 to 2006 (http://www.moh.gov.cn/open/2007tjts/P49.htm). In this study, TB patients, BCG-vaccinated individuals and healthy controls live in many areas with different genetic background. The impacts on sensitivity imposed by the genetic background and regions have been observed (*28*, *29*). Even the same recombinant Rv0934 antigen showed 22% and 28% frequency recognition respectively when tested with Danish TB patients and African population. Besides, the immune response also reportedly varied with the smear status and disease manifestation (*30*, *31*). So the different stage of disease on patients from high and low endemicity regions may also be the reasons resulting in the different sensitivity of the same antigen.

Considering the great heterogeneity from individual to individual and among various regions, we increased the sensitivity with two-protein cocktail while maintaining high specificity. One protein, genetically fusing Rv0934 and Rv0251c together, has 70% sensitivity and 96% specificity. Currently, more proteins combined in various cocktail kits are under test in order to find ideal kits with high sensitivity and specificity.

Conclusion

By comparing the genome sequences of H37Rv and BCG and predicting the antigenicities of the candidate genes, we systemically evaluated 11 purified antigens with a cohort composed of 58 TB patients, 8 healthy volunteers and 50 PPD-negative individuals before and after BCG vaccination. These 11 antigens had at least 92% specificity in healthy controls and six seroantigens (Rv0251c, Rv1973, Rv2376c, Rv2537c, Rv2785c and Rv3873A) were novel. The approach in this study combining comparative genomics, bioinformatics and ELISA techniques will greatly facilitate the identification of novel seroantigens distinguishing *M. tuberculosis* infection from BCG vaccination.

Materials and Methods

Sample population

Serum samples were collected from three different groups of individuals recruited at the Second Hospital Affiliated to the General Hospital of the People's Liberation Army, China. The first group included 58 TB patients (34 males and 24 females) confirmed with the following methods: X-ray, clinical symptoms and patient history. Among the 58 TB patients, 8 had extrapulmonary TB and the other 50 had pulmonary TB. The second group, which comprised 8 healthy volunteers with unknown exposure to TB, confirmed by X-ray and/or clinical findings and patient history, was used as healthy controls in the first round of screening step. They were not checked by skin test with PPD. The third group consisted of 50 PPD-negative individuals checked by skin test. After confirming their negative status, their sera were collected as "before BCG vaccination" and then they were BCG-vaccinated. The sera "after BCG vaccination" were collected three months after vaccination. During the collection of serum, we fully adhered to ethical requirements.

Antibody detection by ELISA

All samples with serum dilution of 1:50 were run in duplicate. The ELISA procedure was performed as follows: 96-well polystyrene microtiter plates were coated for 2 h at 37°C with 100 μL of antigen solution (1 μg/mL for single antigen and 10 μg/mL for PPD) in phosphate-buffered saline (PBS, pH 7.4). The coated plates were washed twice with PBS-T (PBS, 0.05% Tween 20, pH 7.4), and blocked with 2% (weight/volume) bovine serum albumin (BSA) in PBS-T overnight at 4°C. In the next day, serum samples diluted 1:50 in PBS (pH 7.4) were applied and incubated for 1 h at 37°C. The plates were washed three times with PBS-T and then incubated for 1 h with rabbit anti-human IgG antibody conjugated with horseradish peroxidase diluted 1:10,000 in PBS-T. After 3 washings with PBS-T, the enzyme activity was assayed by incubation for 15 min at 37°C with 100 μL of mixtures of tetramethylbenzidine (TMB) (10 μg TMB) per well. 50 μL of 8 N sulfuric acid was added to stop the reaction and the optical density (OD) was measured at 405 nm.

Data analysis

The ELISA results were analyzed by using cutoff values defined by the following equation: Cutoff = OD_{mean} + 2.576×SD, where OD_{mean} and SD stand for the mean OD value and standard deviation for the healthy controls. If the mean OD value for any serum of healthy control is greater than Cutoff, the cutoff value will be recalculated in the above equation after that OD reading is discarded. At the same time, we considered the serum as positive. For statistical analysis, the difference between groups of TB patients and BCG-vaccinated individuals or healthy controls was calculated by the nonparametric Mann-Whitney rank sum test, while the difference between BCG-vaccinated individuals and healthy controls was calculated by Wilcoxon Signed Ranks Test, using the SigmaStat software package (SPSS Inc., Chicago, USA). Differences were considered statistically significant if the P value was below 0.05.

Comparing the identities of the proteins encoded within the genomes of H37Rv and BCG_P

The reference sequences of all proteins in *M. tuberculosis* H37Rv were compared against those of the translated genomes of BCG_P (*M. bovis* BCG strain Pasteur 1173P2) in local server with tBLASTn program. All the reference sequences of proteins in H37Rv and the sequence of BCG_P genome were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/genomes).

Selecting and predicting the antigenecity of the primary candidate proteins

Firstly, we picked out the proteins with low identity (below 60%) between H37Rv and BCG_P, which fell into one of the following groups: (1) PE/PPE or MCE protein family; (2) the assumable secreted or transmembrane proteins. Secondly, reportedly promising antigens or proteins presumably being drastically regulated under different stimulus (such as HSPs) were also selected. Finally, we cloned the candidate proteins with the high antigenic index (above 3.0) predicted by BEPITOPE software (*32*) with the default parameters.

Cloning

Gene-specific primers and annealing temperatures used for PCR and cloning were listed in **Table S1**. PCR was conducted in a 50 μL mixture with 0.25 μL PyroBest DNA polymerase (Takara, Japan), 10 ng of H37Rv genome DNA, 0.4 μM primers, and 250 μM dNTPs. PCR was initiated by a denaturation step at 94°C for 5 min, followed by 25 cycles of denaturation (94°C, 30 s), annealing (temperatures listed in Table S1, 30 s), extension (72°C, 2 min), and a final extension step (72°C, 10 min). The PCR products were cloned into pET28b$^+$ bacterial expression plasmid (Novagen, Germany) at the restriction sites (Table S1). All DNA segments synthesized by PCR were verified by sequencing.

Expression and purification of recombinant proteins

E. coli strain BL21-Gold (DE3) (Stratagene, USA) was used as the cloning host and for the overexpression of all recombinant proteins except Rv1984c, Rv2537c and Rv2785c proteins, which were expressed in *E. coli* strain Rosetta (DE3) (Novagen). Expression was induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) for 3–5 h at 37°C. Bacterial pellets were harvested by centrifugation at 6,000 rpm and 4°C for 10 min, washed in Wash Buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.0). Pellets were thawed and resuspended in Buffer A (20 mM Tris-HCl, 0.5 M NaCl, 5 mM imidazole, pH 8.0), and lysed using a JY92-II Ultrasonic Crasher (Ningbo Scientz Biotechnology Co. Ltd., China) with 60 rounds of 5-second sonication/10-second ice incubation. All the following centrifugation procedures were performed at 12,000 rpm and 4°C for 20 min.

Soluble Rv0251c, Rv2031c, Rv2185c and Rv2376c proteins were purified in the following way: *E. coli* cells overexpressing recombinant proteins were lysed by sonication in Buffer A. The supernatants from the mixtures were loaded on a Ni-NTA column and the bound proteins were eluted stepwise with Buffer A containing from 5 to 300 mM imidazole.

The rest recombinant proteins Rv1971, Rv1973, Rv2537c, Rv2785c, Rv2878c, Rv1563cB, Rv3873B, Rv3804c, Rv1984c, Rv0934, Rv1563cA, Rv1966 and Rv3873A were purified as follows: *E. coli* cells with high levels of recombinant proteins were sonicated in Wash Buffer. The lysates were centrifuged and the supernatants were decanted. The left pellets were resuspended with Buffer B (50 mM Tris-HCl, 5% Triton X-100, 50 mM NaCl, 5 mM EDTA, pH 8.0) and centrifuged, repeated for twice. The washed pellets (inclusion body) were lysed in Buffer C (20 mM Tris-HCl, 0.5 M NaCl, 8 M Urea, pH 8.0) and centrifuged. The supernatants from mixtures were loaded on a Ni-NTA column (Qiagen, Germany), and the bound proteins were eluted stepwise with Buffer C containing from 5 to 200 mM imidazole.

The concentrations of all proteins were determined by the protein A280 method of ND-1000 Spectrophotometer (NanoDrop Technologies, USA).

Acknowledgements

We thank Tao Feng (BIG, CAS) for his technical assistance in bioinformatics analysis. This work was supported by a grant from Chinese Academy of Sciences.

Authors' contributions

LZ and GL collected serum samples. GZ, MZ, LP and JH performed experiments. LZ, GZ and FW conducted data analyses. GZ and SH prepared the manuscript. JY and SH supervised the project and co-wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

1. World Health Organization. 2009. *Global Tuberculosis Control: Epidemiology, Strategy, Financing*, pp.7, 12. WHO Press, Geneva, Switzerland.
2. Hussain, R., *et al.* 1997. Immune response profile in patients with active tuberculosis in a BCG vaccinated area. *Southeast Asian J. Trop. Med. Public Health* 28: 764-773.
3. Lyashchenko, K., *et al.* 1998. Heterogeneous antibody responses in tuberculosis. *Infect. Immun.* 66: 3936-3940.
4. Houghton, R.L., *et al.* 2002. Use of multiepitope polyproteins in serodiagnosis of active tuberculosis. *Clin. Diagn. Lab. Immunol.* 9: 883-891.
5. Weldingh, K., *et al.* 2005. Assessing the serodiagnostic potential of 35 *Mycobacterium tuberculosis* proteins

and identification of four novel serological antigens. *J. Clin. Microbiol.* 43: 57-65.

6. Lodes, M.J., *et al.* 2001. Serological expression cloning and immunological evaluation of MTB48, a novel *Mycobacterium tuberculosis* antigen. *J. Clin. Microbiol.* 39: 2485-2493.
7. Brusasca, P.N., *et al.* 2001. Immunological characterization of antigens encoded by the RD1 region of the *Mycobacterium tuberculosis* genome. *Scand. J. Immunol.* 54: 448-452.
8. Vordermeier, H.M., *et al.* 2001. Use of synthetic peptides derived from the antigens ESAT-6 and CFP-10 for differential diagnosis of bovine tuberculosis in cattle. *Clin. Diagn. Lab. Immunol.* 8: 571-578.
9. Cockle, P.J., *et al.* 2002. Identification of novel *Mycobacterium tuberculosis* antigens with potential as diagnostic reagents or subunit vaccine candidates by comparative genomics. *Infect. Immun.* 70: 6996-7003.
10. He, X.Y., *et al.* 2003. Comparative proteome analysis of culture supernatant proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra. *Microbes Infect.* 5: 851-856.
11. Jungblut, P.R., *et al.* 1999. Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. *Mol. Microbiol.* 33: 1103-1117.
12. Mattow, J., *et al.* 2003. Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. *Electrophoresis* 24: 3405-3420.
13. Chakhaiyar, P., *et al.* 2004. Regions of high antigenicity within the hypothetical PPE major polymorphic tandem repeat open-reading frame, Rv2608, show a differential humoral response and a low T cell response in various categories of patients with tuberculosis. *J. Infect. Dis.* 190: 1237-1244.
14. Choudhary, R.K., *et al.* 2003. PPE antigen Rv2430c of *Mycobacterium tuberculosis* induces a strong B-cell response. *Infect. Immun.* 71: 6338-6343.
15. Delogu, G. and Brennan, M.J. 2001. Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis*. *Infect. Immun.* 69: 5606-5611.
16. Demangel, C., *et al.* 2004. Cell envelope protein PPE68 contributes to *Mycobacterium tuberculosis* RD1 immunogenicity independently of a 10-kilodalton culture filtrate protein and ESAT-6. *Infect. Immun.* 72: 2170-2176.
17. Okkels, L.M., *et al.* 2003. PPE protein (Rv3873) from DNA segment RD1 of *Mycobacterium tuberculosis*: strong recognition of both specific T-cell epitopes and epitopes conserved within the PPE family. *Infect. Immun.* 71: 6116-6123.
18. Fleischmann, R.D., *et al.* 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* 184: 5479-5490.
19. Cole, S.T., *et al.* 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393: 537-544.
20. Yuan, Y., *et al.* 1998. The 16-kDa alpha-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. *Proc. Natl. Acad. Sci. USA* 95: 9578-9583.
21. Sherman, D.R., *et al.* 2001. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. *Proc. Natl. Acad. Sci. USA* 98: 7534-7539.
22. Boon, C. and Dick, T. 2002. *Mycobacterium bovis* BCG response regulator essential for hypoxic dormancy. *J. Bacteriol.* 184: 6760-6767.
23. Andersen, P. and Doherty, T.M. 2005. The success and failure of BCG—implications for a novel tuberculosis vaccine. *Nat. Rev. Microbiol.* 3: 656-662.
24. Launois, P., *et al.* 1994. T-cell-epitope mapping of the major secreted mycobacterial antigen Ag85A in tuberculosis and leprosy. *Infect. Immun.* 62: 3679-3687.
25. McShane, H., *et al.* 2002. Protective immunity against *Mycobacterium tuberculosis* induced by dendritic cells pulsed with both $CD8^+$- and $CD4^+$-T-cell epitopes from antigen 85A. *Infect. Immun.* 70: 1623-1626.
26. Espitia, C., *et al.* 1989. A 38-kD *Mycobacterium tuberculosis* antigen associated with infection. Its isolation and serologic evaluation. *Clin. Exp. Immunol.* 77: 373-377.
27. Rosenkrands, I., *et al.* 1999. Differential T-cell recognition of native and recombinant *Mycobacterium tuberculosis* GroES. *Infect. Immun.* 67: 5552-5558.
28. Huygen, K., *et al.* 1993. Influence of genes from the major histocompatibility complex on the antibody repertoire against culture filtrate antigens in mice infected with live *Mycobacterium bovis* BCG. *Infect. Immun.* 61: 2687-2693.
29. Wheeler, E., *et al.* 2006. Genome-wide scan for loci influencing quantitative immune response traits in the Belem family study: comparison of methods and summary of results. *Ann. Hum. Genet.* 70: 78-97.
30. Davidow, A., *et al.* 2005. Antibody profiles characteristic of *Mycobacterium tuberculosis* infection state. *Infect. Immun.* 73: 6846-6851.
31. Imaz, M.S. and Zerbini, E. 2000. Antibody response to culture filtrate antigens of *Mycobacterium tuberculosis* during and after treatment of tuberculosis patients. *Int. J. Tuberc. Lung Dis.* 4: 562-569.
32. Odorico, M. and Pellequer, J.L. 2003. BEPITOPE: predicting the location of continuous epitopes and patterns in proteins. *J. Mol. Recognit.* 16: 20-22.

33. Wilkinson, K.A., *et al.* 2005. Infection biology of a novel alpha-crystallin of *Mycobacterium tuberculosis*: Acr2. *J. Immunol.* 174: 4237-4243.
34. Andersen, A.B. and Hansen, E.B. 1989. Structure and mapping of antigenic domains of protein antigen b, a 38,000-molecular-weight protein of *Mycobacterium tuberculosis*. *Infect. Immun.* 57: 2481-2488.
35. De Smet, K.A., *et al.* 2000. Three pathways for trehalose biosynthesis in mycobacteria. *Microbiology* 146: 199-208.
36. Ahmad, S., *et al.* 2004. Mammalian cell-entry proteins encoded by the mce3 operon of *Mycobacterium tuberculosis* are expressed during natural infection in humans. *Scand. J. Immunol.* 60: 382-391.
37. Weldingh, K., *et al.* 1998. Two-dimensional electrophoresis for analysis of *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. *Infect. Immun.* 66: 3492-3500.
38. Rosenkrands, I., *et al.* 2000. Towards the proteome of *Mycobacterium tuberculosis*. *Electrophoresis* 21: 3740-3756.
39. Matsumoto, S., *et al.* 1995. Cloning and sequencing of a unique antigen MPT70 from *Mycobacterium tuberculosis* H37Rv and expression in BCG using *E. coli*-mycobacteria shuttle vector. *Scand. J. Immunol.* 41: 281-287.
40. Webb, J.R., *et al.* 1998. Molecular cloning, expression, and immunogenicity of MTB12, a novel low-molecular-weight antigen secreted by *Mycobacterium tuberculosis*. *Infect. Immun.* 66: 4208-4214.
41. Nagai, S., *et al.* 1991. Isolation and partial characterization of major protein antigens in the culture fluid of *Mycobacterium tuberculosis*. *Infect. Immun.* 59: 372-382.
42. Turneer, M., *et al.* 1988. Humoral immune response in human tuberculosis: immunoglobulins G, A, and M directed against the purified P32 protein antigen of *Mycobacterium bovis* bacillus Calmette-Guerin. *J. Clin. Microbiol.* 26: 1714-1719.
43. Khan, I.H., *et al.* 2008. Profiling antibodies to *Mycobacterium tuberculosis* by multiplex microbead suspension arrays for serodiagnosis of tuberculosis. *Clin. Vaccine Immunol.* 15: 433-438.
44. Johnson, S., *et al.* 2001. Characterization of the secreted MPT53 antigen of *Mycobacterium tuberculosis*. *Infect. Immun.* 69: 5936-5939.
45. Sanchez-Rodriguez, C., *et al.* 2002. An IgG antibody response to the antigen 85 complex is associated with good outcome in Mexican Totonaca Indians with pulmonary tuberculosis. *Int. J. Tuberc. Lung Dis.* 6: 706-712.

Supporting Online Material

Table S1

DOI: 10.1016/S1672-0229(08)60039-X