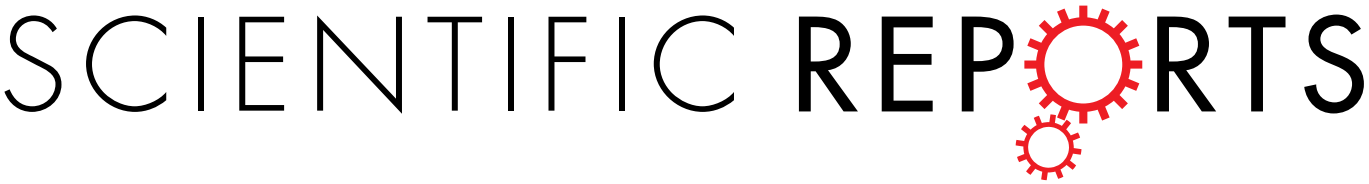

OPEN

Evaluation of 5 Novel protein biomarkers for the rapid diagnosis of pulmonary and extra-pulmonary tuberculosis: preliminary results

Received: 12 July 2016
Accepted: 06 February 2017
Published: 24 March 2017

Amit Singh[1,*], Anil Kumar Gupta[1,*], Krishnamoorthy Gopinath[1,†], Pawan Sharma[2,‡] & Sarman Singh[1]

Improved methods are required for the early and accurate diagnosis of tuberculosis, especially in the patients with smear-negative disease. Several biomarkers have been tried but most have shown poor sensitivity or specificity. In present study we aimed to evaluate the diagnostic utility of five novel antigens identified earlier by us. This is an initial study conducted on 250 subjects. The five recombinant antigens, named as rSS1 (Rv2145c), rSS2 (Rv0164), rSS3 (Rv1437), rSS4 (Rv1827) and rSS5 (Rv2970c), were expressed in pQE-30 expression vector, purified and their sero-diagnostic efficacy was evaluated in an unblinded manner using dot-blot and ELISA methods. The sensitivity and specificity of these novel antigens were compared with commercially available standard esat6 and 38 kDa antigens. Bacteriologically confirmed TB patients, non-TB disease controls and healthy individuals were included. which are based on novel antigen or novel technology, Area under curve (AUC) of the selected antigens were 0.98 (0.98–0.99) for rSS1, 0.88 (0.84–0.92) for rSS2, 0.88 (0.84–0.92) for rSS3, 0.95 (0.93–0.98) for rSS4 and 0.99 (0.98–1.0) for rSS5. Receiver operative characteristic (ROC) curve showed highly significant difference between TB and healthy subjects ($p = <0.001$). These initial findings, show that the recombinant antigens rSS1, rSS4 and rSS5 could be used as highly potential biomarkers for the serological diagnosis of active TB.

Tuberculosis (TB) still remains as a major health problem in the developing countries and is rated as the number one killer infectious disease. World Health Organization (WHO) estimated 9.6 million new cases and more than 1.5 million deaths annually worldwide and India has the world's largest tuberculosis epidemics[1]. The biggest hurdle in the control and timely management of tuberculosis is non-availability of rapid, accurate and cost-effective test. In remote areas, TB diagnostics are dependent on the microscopic observation of acid-fast bacilli (AFB) in the clinical samples or by bacteriological culture analysis. Although, AFB-smear staining allows rapid and cost-effective diagnosis of tuberculosis but it has comparatively low sensitivity, especially in children and immuno-compromised patients. The current nucleic acid amplification based tests (NAAT) such as Polymerase Chain Reaction (PCR); Xpert MTB/Rif assay; and Line Probe Assay (LPA) are rapid and sensitive but are expensive. A recently held international TB consortium meeting at New Delhi, outlined the urgent need for newer and better point-of care tests.

The serological tests have been an attractive diagnostic tool due to their convenience, rapidity and easy implementation in the national programmes. These tests have contributed significant role in the early diagnosis and management of several infectious diseases including Human Immunodeficiency Virus (HIV), hepatitis A, B, C, E, Leishmaniasis, malaria etc. However, previous attempts to diagnose TB by serology have met with limited success due to low sensitivity and specificity[2,3]. Hence, WHO banned these serological tests in 2011 and later on in 2012 Government of India also banned import manufacturing and sale of these kits[3,4]. Therefore rapid, inexpensive and accurate antibody based tests for TB diagnosis are urgently required[5]. Earlier, we had reported

[1]Division of Clinical Microbiology & Molecular Medicine, All India Institute of Medical Sciences, New Delhi, India. [2]International Centre for Genetic Engineering and Biotechnology, New Delhi, India. [†]Present address: Max Planck Institute for Infection Biology, Charitéplatz, Berlin, Germany. [‡]Present address: North East Region-BPMC, Department of Biotechnology, Government of India, New Delhi, India. [*]These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.S. (email: sarman_singh@yahoo.com)

S. No.	Gene(s)	Plasmid	Cloning vector (cloning sites)	Primers used for amplification
1	rSS1	pQE30	pQE-SS1	F 5′-CGGGATCCATGCCGCTTACACCTGCC-3′
			(*BamH*I & *Hind*III)	R 5′-CCCAAGCTTCTAGTTTTTGCCCCGGTTGAA-3′
2	rSS2	pQE30	pQE-SS2	F 5′-CGGGATCCATGACGGCAATCTCGTGCTC-3′
			(*BamH*I & *Hind*III)	R 5′-CCCAAGCTTTTAGCTGGCCGCCAGCTG-3′
3	rSS3	pQE30	pQE-SS3	F 5′-CGGGATCCATGAGCGTTGCAAACCTCAAG-3′
			(*BamH*I & *Hind*III)	R 5′-CCCAAGCTTTCACAAAACTCCTCCGGTTGG-3′
4	rSS4	pQE30	pQE-SS4	F 5′-CGGGATCCGTG ACG GAC ATG AAC CCG GA-3′
			(*BamH*I & *Hind*III)	R 5′-CCCAAGCTTTCA CGG GCC CCC GGT ACT-3′
5	rSS5	pQE30	pQE30-SS5	F 5′-CGGGATCCATGACCAAGAGTCTGCCAGG-3
			(*BamH*I & *Hind*III)	R 5′-CCCAAGCTT#TCAAACCCGGCTAAGGTGC-3

Table 1. Vectors and primers used for cloning of the five novel proteins of *Mycobacterium tuberculosis* (Mtb) named as rSS1, rSS2, rSS3, rSS4 and rSS5. Nucleotide sequences recognised by the restriction enzyme used for cloning are underlined.

few differentially expressed proteins that were over expressed during the active disease and development of drug resistance *in-vivo*[6]. Through bioinformatics analysis, we selected five potential proteins based on their antigenicity. In the present study, we further evaluated these antigens for their diagnostic potential by dot-blot and enzyme linked immuno-sorbent assay (ELISA) in various patient and control groups.

Results

Cloning, expression and purification of mycobacterial antigens. The five novel proteins as described in the material and method section have been named as rSS1 (Rv2145c), rSS2 (Rv0164), rSS3 (Rv1437), rSS4 (Rv1827) and rSS5 (Rv2970c) and a patent has been filed (1752/DEL/2008). The nucleotide sequences of the amplicons were analysed and submitted to GenBank under accession number KC147003, KC147004, KC147005, KC147006 and KC147008, respectively. To generate recombinant antigens, PCR amplification of target genes from genomic DNA of the *Mtb* clinical isolates[7], was carried out using gene specific primers (Table 1). All the selected genes were successfully amplified with product size of 783 bp (rSS1), 458 bp (rSS2), 1239 bp (rSS3), 489 bp (rSS4) and 1131 bp (rSS5), respectively with appropriate restriction sites (Fig. 1A). The products were cloned and desired proteins were expressed in expression vector as detailed in method section. All recombinant proteins were purified by Ni^{2+}-NTA affinity chromatography under denaturing conditions from inclusion bodies except the rSS4 protein which was purified under native condition. The purity (>96%) of N-terminal His-tagged recombinant proteins was analysed by SDS-PAGE (Fig. 1B). The observed molecular weights of rSS1, rSS2, rSS3, rSS4 and rSS5 proteins were approximately 28 kDa, 18 kDa, 42 kDa, 17 kDa and 42 kDa, respectively. The yield of purified recombinant proteins were 3 g (rSS1), 0.7 g (rSS2), 1 g (rSS3), 0.3 g (rSS4) and 0.28 g (rSS5) per litre of the bacterial culture.

Detection of antibody response against five antigens by immunoblot assay. The immune reactivity of five individual proteins was optimized and used to check diagnostic potential in serum samples of different patient groups using dot-blot assay. Two standard proteins (esat6 and 38 kDa) were kind gift from Prof. VK Chaudhary. The minimum detection concentrations of the purified antigens were found to be 5 ng (rSS1), 10 ng (rSS2), 12.5 ng (rSS3), 10 ng (rSS4), 10 ng (rSS5), 10 ng (38 kDa Ag) and 10 ng (esat6 Ag) of the antigens at 1:200 serum dilution in dot-blot assay. The clinical and laboratory characteristics of the TB patients and healthy controls (HC) included in the study are summarized in Table 2. A total of 250 subjects were included in this study. Of these, 140 were culture confirmed tuberculosis patients [111 pulmonary tuberculosis (PTB) and 29 extra-pulmonary tuberculosis (EPTB)] and 110 controls [60 non-tuberculosis-diseased controls (DC) and 50 healthy controls (HC)]. Of the 111 PTB patients, 15 were HIV-positive and 96 HIV-negative. Among the EPTB patients, five were HIV positive and 15 HIV-negative. Of the 250 subjects 81 (32.4%) had Bacillus Calmette-Guérin (BCG) vaccination with discernible scar and 169 (67.6%) had no vaccination/no discernible scar. Eighty patients (57.1%) were Mantoux/tuberculin skin test (TST) positive positive, while 170 (42.9%) were Mantoux test negative. The control sera (DC plus HC) were included for the assessment of the specificity of selected five proteins in ELISA and dot-blots assay. Of which, 60 subjects had diseases other than tuberculosis, while 50 were HC (with no known past history of TB and TST negative). Eighteen (30%) DC and 14 (28%) HC gave history of *M. bovis* BCG vaccination and were verified by scar. The sera were tested parallally using each recombinant antigen and the results are shown in Tables (3 and 4).

Dot-blot screening of recombinant antigens for TB diagnosis using dot-blot assay. The antibody response to the novel antigens was visually analysed (Fig. 2) and results were compared with gold standard MGIT™ 960 culture for sensitivity calculation. The recombinant esat6 and 38 kDa Ag proteins were used as reference antigens. The sensitivity and specificity of esat6 and 38 kDa antigens ranged between 72.4% to 94% and 54.5% to 66.4% respectively. The pooled sensitivity of 38 kDa Ag in PTB, EPTB, MDR-TB cases was 85.6%, 86.2% 94%, while that of esat6 was 85.6%, 72.4%, and 90% respectively. However, in comparison to the reference antigens, the overall performance of our five antigens was much superior. The rSS1, rSS2, rSS3, rSS4, rSS5 antigens showed superior sensitivity and specificity both ranging from 86.2% to 99.1% and 89.1% to 100%, respectively.

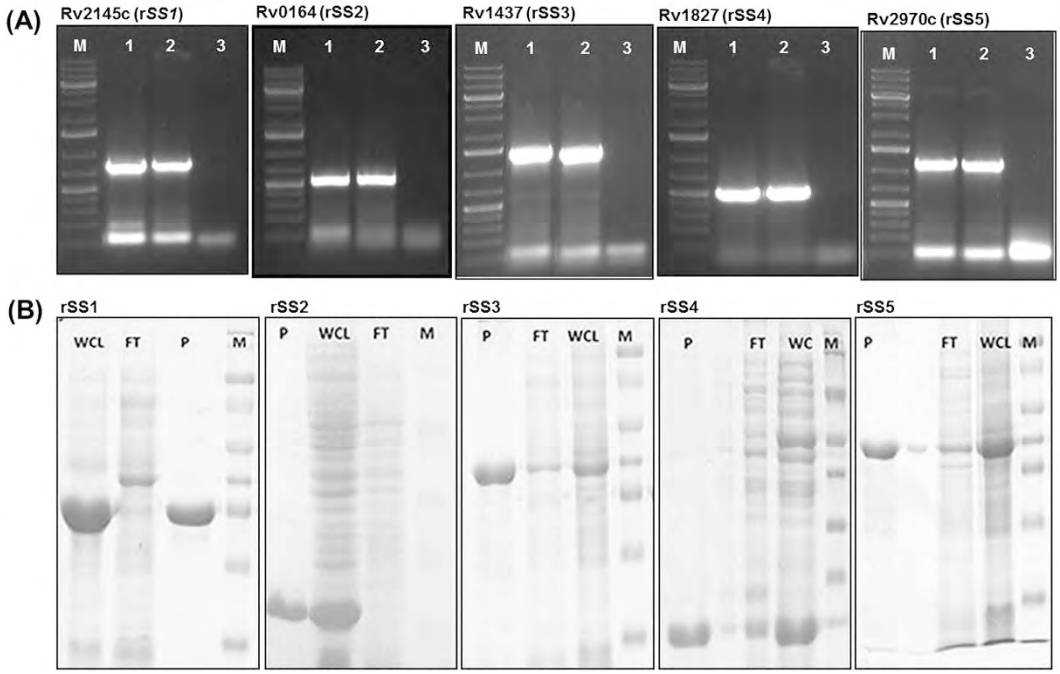

Figure 1. Gene amplification, cloning, expression and purification . (**A**) Agarose gel electrophoresis for PCR products. Lane M, 1 kb DNA Molecular-size marker; lane 1, 2: amplification of respective genes -rSS1 (783 bp), rSS2 (468 bp), rSS3 (489 bp), rSS4 (1239 bp), and rSS5 (1131 bp) from *M. tuberculosis* clinical isolates. lane 3 is negative control. (**B**) SDS-PAGE analyses of purified recombinant proteins (rSS1, rSS2, SS3, rSS4 and rSS5). Proteins were visualized with Coomassie brilliant blue staining. Abbreviations: WCL: Whole Cell Lysate, FT: Flow Through, P: Purified Protein, M: Marker (Protein).

The detailed sensitivity and specificity values of each antigen are given in Table (3). In PTB cases, the sensitivity of these antigens ranged between 99.1% and 93.7% and specificity in various control groups was between 89.1% and 100%. Interestingly, the rSS1, rSS4 and rSS5 antigens showed 100% sensitivity in HIV-PTB cases (see Supplementary Table 1). In EPTB cases also, the pooled sensitivity and specificity were very high (Table 3). Antigens rSS1, rSS2, rSS4 and rSS5 showed 100% sensitivity in HIV-EPTB cases and smear positive EPTB cases (Supplementary Table 1). In Mantoux positive TB cases, 100% (8/8) sensitivity was observed for rSS5, rSS4 and rSS1 antigens followed by 87.5% for rSS2, rSS3, Ag38 kDa and esat6 antigens (Supplementary Table 2). In smear negative culture positive cases, antigen rSS5 showed maximum sensitivity of 96.0%. The sensitivity and specificity rates of the 5 antigens by dot-blot assay, in MDR-TB cases were 100% while sensitivity of 38 kDa Ag and esat6 Ag were found 94% and 90% respectively (Table 3 and Supplementary Table 3).

ELISA screening of purified recombinant antigens for TB diagnosis using ELISA test. The optimum concentration of recombinant antigen yielding high specificity was determined at 25 ng/well for rSS1 and rSS5, 50 ng/well for rSS2, rSS3, rSS4, 38 kDa Ag, esat6 Ag and the serum dilutions of primary and secondary conjugated antibodies were found to be 1:50 and 1:15000, respectively. Using these dilutions all of the 250 serum samples (Table 2) were analysed by indirect ELISA. The cut-off value (Mean $\pm$ 2 SD) of the ELISA was determined by area under curve (AUC) from 140 TB patient serum samples and 110 HC and DC serum samples (Fig. 3). The cut-off values were 0.597 for rSS1, 0.489 for rSS2, 0.540 for rSS3, 0.411 for rSS4, 0.410 for rSS5, 0.57 for 38 kDa Ag and 0.56 for esat6 Ag. The AUC were 0.98 (0.98–0.99) for rSS1, 0.88 (0.84–0.92 for rSS2, 0.88 (0.84–0.92) for rSS3, 0.95 (0.93–0.98) for rSS4, 0.99 (0.98–1.0) for rSS5 0.90(0.88–0.92) for 38 kDa Ag and 0.81(0.79–0.83) for esat6 Ag (Fig. 4).

The highest sensitivity and specificity values for these recombinant antigens were 96.5% & 98.2% for rSS5, 96.4% & 97.3% for rSS1, 73% & 87.3% for rSS2, 75.9% & 86.4% for rSS3, 91.9% and 94.5% for rSS4, respectively (Table 4). The overall sensitivity of esat6 & 38 kDa Ag was 81.3% and 91.25%, while specificity was 70.9% and 60.9% respectively. There was a clear difference in the antibody levels observed in healthy controls *vs* TB patients ($P_{rSS5} < 0.001$, $P_{rSS1} < 0.001$, $P_{rSS4} < 0.001$, $P_{rSS2} < 0.001$ and $P_{r}SS3 < 0.001$, respectively) (Table 4; Fig. 4), However, rSS5, rSS1 and rSS4 antigens showed very good activity over other antigens in MDR-TB cases. The pooled sensitivity rates in PTB cases by ELISA for antigens rSS1, rSS2, rSS3, rSS4 and rSS5 were between 73.0% to 96.4%, (Supplementary Table 4). In EPTB cases, the highest sensitivity (96.5%) and specificity (98.2%) was shown by rSS5, followed by 89.7% & 97.3% for rSS1, 89.7% & 94.5% for rSS4, 89.7% & 70.9% for esat6 Ag, 89.7% & 60.9% for 38 kDa Ag, 75.9% & 86.7% for rSS3, 69.0% & 87.3% for rSS2 (Fig. 4 and Supplementary Table 5). The sensitivity and specificity of five antigens in the MDR-TB patients are given in Supplementary Table (6). In comparison to reference antigens, our three antigens (rSS1, rSS4, rSS5) showed extraordinary performance in

S. No.	Category (n)	Mean Age (Yr) ± SD	Gender		HIV status		BCG Vaccination		Mantoux test		Smear		MGIT Culture (%)	MDR (%)
			Male (%)	Female (%)	Pos (%)	Neg (%)	Yes (%)	No/uk (%)	Pos (%)	Neg/uk (%)	Pos (%)	Neg (%)		
1	PTB (n = 111)	35.0 ± 15.2	63 (56.8)	48 (43.2)	15 (13.5)	96 (86.5)	42 (37.8)	69 (62.2)	68 (61.3)	43 (38.7)	72 (64.9)	39 (35.1)	111 (100)	46 (41.4)
2	EPTB (n = 29)	31.0 ± 16.1	15 (51.7)	14 (48.3)	5 (17.2)	24 (82.8)	10 (34.5)	19 (65.5)	11 (37.9)	18 (62.1)	4 (13.8)	25 (86.2)	29 (100)	4 (13.8)
4	HC (n = 50)	30.0 ± 13.2	30 (60.0)	20 (40.0)	0	50 (100)	14 (28.0)	36 (72.0)	—	50 (100)	0	50 (100)	0	0
5	DC (non-TB) n = 60	31.4 ± 14.6	34 (56.7)	26 (43.3)	20 (33.3)	40 (66.7)	18 (30.0)	42 (70.0)	—	60 (100)	0	60 (100)	0	0
Total	**250**	**31.8 ± 14.8**	**142 (56.8)**	**108 (43.2)**	**40 (16.0)**	**210 (84.0)**	**84 (33.6)**	**166 (66.4)**	**79 (31.6)**	**171 (68.4)**	**76 (30.4)**	**174 (69.6)**	**140 (56.0)**	**50 (20.0)**

Table 2. Detailed clinical and demographical profile of subjects (N = 250) included in the study.

	Sensitivity [n (%; 95% CI)]	Specificity [n (%; 95% CI)]	PPV (%)	NPV (%)	LRP (95% CI)	DA %, (95% CI)
PTB cases (n = 111)						
rSS5 (Rv2970c)	*110/111 (99.1%; 95.1, 99.8)*	*108/110 (98.2%; 93.6, 99.6)*	*98.2%*	*99.1%*	54.5 (20.4–145.2)	98.6% (96.1, 99.5)
rSS1 (Rv2145c)	*110/111 (99.1%; 95.1, 99.8)*	*110/110 (100%; 96.6, 100)*	*100%*	*99.1%*	—	*99.5% (97.5, 99.9)*
rSS4 (Rv1827)	*109/111 (98.2%; 93.7, 99.5)*	*103/110 (93.6%; 87.4, 96.9)*	*94.0%*	*98.1%*	*15.4 (11.6–20.4)*	*95.9% (92.4, 97.8)*
rSS2 (Rv0164)	*107/111 (96.4%; 91.1, 98.6)*	*102/110 (92.7%; 86.3, 96.3)*	*93.0%*	*96.2%*	*13.2 (10.4–16.9)*	*94.6% (90.7, 96.9)*
rSS3 (Rv1437)	*104/111 (93.7%; 87.5, 96.9)*	*98/110 (89.1%; 81.9, 93.6)*	*89.7%*	*93.3%*	*8.6 (7.3–10.1)*	*91.4% (87.0, 94.4)*
Esat6 Ag	*95/111 (85.6%; 77.9, 90.9)*	*73/110 (66.4%; 57.1, 74.51)*	*72.0%*	*82.0%*	*2.5 (2.4–2.7)*	*76.0% (70.0, 81.2)*
38 kDa Ag	*95/11116 (85.6%; 77.9, 90.9)*	*60/110 (54.5%; 45.2, 63.5)*	*65.5%*	*78.9%*	*1.9 (1.8–2.0)*	*70.1% (63.8, 75.8)*
EPTB cases (n = 29)						
rSS5 (Rv2970c)	*28/29 (96.5%; 82.8, 99.4)*	*108/110 (98.2%; 93.6, 99.6)*	*93.3%*	*99.1%*	*53.1 (19.9–141.8)*	*97.8% (93.8, 99.3)*
rSS1 (Rv2145c)	*27/29 (93.1%; 78.0, 98.1)*	*110/110 (100%; 96.6, 100)*	*100%*	*98.2%*	—	*98.6% (94.9, 99.6)*
rSS4 (Rv1827)	*27/29 (93.1%; 78.0, 98.1)*	*103/110 (93.6%; 87.4, 96.9)*	*79.4%*	*98.1%*	*14.6 (11.0–19.5)*	*93.5% (88.1, 96.6)*
rSS2 (Rv0164)	*25/29 (86.2%; 69.4, 94.5)*	*102/110 (92.7%; 86.3, 96.3)*	*75.8%*	*96.2%*	*11.8 (9.2–15.3)*	*91.4% (85.5, 95.0)*
rSS3 (Rv1437)	*26/29 (89.7%; 73.6, 96.4)*	*98/110 (89.1%; 81.9, 93.6)*	*68.4%*	*97.0%*	*8.2 (6.9–9.8)*	*89.2% (83.0, 93.3)*
Esat6 Ag	*21/29 (72.4%; 54.3, 85.3)*	*73/110 (66.4%; 57.1, 74.51)*	*36.2%*	*90.1%*	*2.1 (2.0–2.3)*	*67.6 (59.5–74.8)*
38 kDa Ag	*25/29 (86.2%; 69.4, 94.5)*	*60/110 (54.5%; 45.2, 63.5)*	*33.4%*	*93.7%*	*2.0 (1.8–2.0)*	*61.1% (52.8, 68.8)*
MDR-TB Cases [(n = 50)						
rSS5 (Rv2970c)	*50/50 (100%; 92.9, 100)*	*108/110 (98.2%; 93.6, 99.6)*	*96.1%*	*100%*	*55.0 (20.6–146.5)*	*98.7% (95.6, 99.7)*
rSS1 (Rv2145c)	*50/50 (100%; 92.9, 100)*	*110/110 (100%; 96.6, 100)*	*100%*	*100%*	—	*100% (97.7, 100)*
rSS4 (Rv1827)	*50/50 (100%; 92.9, 100)*	*103/110 (93.6%; 87.4, 96.9)*	*87.7%*	*100%*	*15.7 (11.9–20.8)*	*95.6% (91.2, 97.9)*
rSS2 (Rv0164)	*50/50 (100%; 92.9, 100)*	*102/110 (92.7%; 86.3, 96.3)*	*86.2%*	*100%*	*13.7 (10.8–17.6)*	*95.0% (90.4, 97.4)*
rSS3 (Rv1437)	*50/50 (100%; 92.9, 100)*	*98/110 (89.1%; 81.9, 93.6)*	*80.6%*	*100%*	*9.2 (7.8–10.8)*	*92.5% (87.3, 95.7)*
Esat6 Ag	*45/50 (90.0%;78.6, 95.6)*	*73/110 (66.4%; 57.1, 74.51)*	*54.9%*	*93.6%*	*2.7 (2.5–2.8)*	*73.7% (66.4, 79.9)*
38 kDa Ag	*47/50 (94.0%;83.8, 97.9)*	*60/110 (54.5%; 45.2, 63.5)*	*48.4%*	*95.2%*	*2.1 (2.0–2.2)*	*66.9% (59.3, 73.7)*

Table 3. Sensitivity and specificity of 5 novel recombinant antigens by dot-blot assay [PTB = 111, EPTB = 29 and controls = 110]. Pos: Positive, Neg: Negative, CI: Confidence interval, PPV: Positive predictive value, NPV: Negative predictive value, LRP: likelihood ratio for positive test, DA: Diagnostic accuracy. *Also see supplementary file 1.

PTB cases (Supplementary Table 4). In EPTB cases, only rSS5 antigen showed very high sensitivity (Fig. 4 and Supplementary Table 5).

Discussion

Rapid and accurate diagnosis of tuberculosis is crucial to facilitate early treatment initiation, reducing disease transmission and preventing emergence of drug resistant strains. The currently used methods are either insensitive, time consuming, costly or require high technical skill and laboratory infrastructure. In search of rapid and cost-effective diagnosis of tuberculosis, serological test have been considered an attractive option. In recent past decades, a variety of serological assays have been developed such as latex agglutination, ELISA, indirect immunofluorescence and rapid immunochromatographic tests[8–11]. The rapid serological tests have successfully been used to combat several infectious diseases like malaria, leishmaniasis, HIV, hepatitis viral infections etc. to name a few. These serological tests, if developed successfully for tuberculosis, will have an additional advantage over all other currently used tests, specially in patients who are unable to produce sputum (such as children and HIV/AIDS patients) and patients with EPTB. These serological tests will also have potential to develop point of care (POC) tests that can reach to lowest levels of health services.

	Sensitivity [n (%; 95% CI)]	Specificity [n (%; 95% CI)]	PPV (%)	NPV (%)	LRP (95% CI)	DA%, (95% CI)
PTB cases (n = 111)						
rSS5 (Rv2970c)	*107/111 (96.4%; 91.1, 98.6)*	*108/110 (98.2%; 93.6, 99.5)*	*98.2%*	*96.4%*	*53.2 (19.9–141.4)*	*97.3% (94.2, 98.7)*
rSS1 (Rv2145c)	*107/111 (96.4%; 91.1, 98.6)*	*107/110 (97.3%; 92.3, 99.1)*	*97.3%*	*96.4%*	*35.3 (18.4–68.0)*	*96.8% (93.6, 98.5)*
rSS4 (Rv1827)	*102/111 (91.9%; 85.3, 95.7)*	*104/110 (94.5%; 88.6, 97.5)*	*94.5%*	*92.0%*	*16.8 (12.1–23.4)*	*93.2% (89.1, 95.8)*
rSS2 (Rv0164)	*81/111 (73.0%; 64.0, 80.4)*	*96/110 (87.3%; 79.8, 92.3)*	*85.3%*	*76.2%*	*5.7 (4.9–6.6)*	*80.1% (74.3, 84.8)*
rSS3 (Rv1437)	*82/111 (73.9%; 65.0, 81.1)*	*95/110 (86.7%; 78.7, 91.6)*	*84.5%*	*76.6%*	*5.4 (4.7–6.2)*	*80.1% (74.3, 84.8)*
Esat6 Ag	*81/111 (72.97%;64.0, 80.4)*	*78/110 (70.9%; 61.8, 78.6)*	*71.7%*	*72.2*	*2.5 (2.3–2.7)*	*71.9% (65.7–77.5)*
38 kDa Ag	*103/11 (92.8%;86.4, 96.3)*	*67/110 (60.9%; 51.6, 69.5)*	*70.5%*	*89.3%*	*2.4 (2.3–2.5)*	*76.9% (70.9, 82.0)*
EPTB cases (n = 29)						
rSS5 (Rv2970c)	*28/29 (96.5%; 82.8, 99.4)*	*108/110 (98.2%; 93.6, 99.5)*	*93.3%*	*99.1%*	*53.1 (19.9–141.8)*	*97.8% (93.8, 99.3)*
rSS1 (Rv2145c)	*26/29 (89.7%; 73.6, 96.4)*	*107/110 (97.3%; 92.3, 99.1)*	*89.7%*	*97.3%*	*32.9 (17.0–63.7)*	*95.7% (90.9, 98.0)*
rSS4 (Rv1827)	*26/29 (89.7%; 73.6, 96.4)*	*104/110(94.5%; 88.6, 97.5)*	*81.2%*	*97.2%*	*16.4 (11.7–23.0)*	*93.5% (88.1, 96.6)*
rSS2 (Rv0164)	*20/29 (69.0%; 50.8, 82.7)*	*96/110 (87.3%; 79.8, 92.3)*	*58.8%*	*91.4%*	*5.4 (4.5–6.5)*	*83.4% (76.4, 88.7)*
rSS3 (Rv1437)	*22/29 (75.9%; 57.9, 87.8)*	*95/110 (86.7%; 78.7, 91.6)*	*59.5%*	*93.1%*	*5.6 (4.7–6.5)*	*84.2% (77.2, 89.3)*
Esat6 Ag	*26/29 (89.7%; 73.6, 96.4)*	*78/110 (70.9%; 61.8, 78.6)*	*44.8%*	*96.3%*	*3.1 (2.9–3.3)*	*74.8% (67.8, 81.3)*
38 kDa Ag	*26/29 (89.7%; 73.6, 96.4)*	*67/110 (60.9%; 51.6, 69.5)*	*37.7%*	*95.7%*	*2.3 (2.2–2.4)*	*66.9% (58.7, 74.2)*
MDR-TB Cases [(n = 50)						
rSS5 (Rv2970c)	*50/50 (100%; 92.9, 100)*	*108/110 (98.2%; 93.6, 99.5)*	*96.1%*	*100%*	*55.0 (20.6–146.5)*	*98.7% (95.6, 99.6)*
rSS1 (Rv2145c)	*50/50 (100%; 92.9, 100)*	*107/110 (97.3%; 92.3, 99.1)*	*94.3%*	*100%*	*37.7 (19.1–70.5)*	*98.1% (94.6, 99.4)*
rSS4 (Rv1827)	*50/50 (100%; 92.9, 100)*	*104/110(94.5%; 88.6, 97.5)*	*89.3%*	*100%*	*18.3 (13.2–25.4)*	*96.2% (92.1, 98.3)*
rSS2 (Rv0164)	*38/50 (76.0%; 62.6, 85.7)*	*96/110 (87.3%; 79.8, 92.3)*	*73.1%*	*88.9%*	*6.0 (5.1–7.0)*	*83.7% (77.2, 88.7)*
rSS3 (Rv1437)	*41/50 (82.0%; 69.2, 90.2)*	*95/110 (86.7%; 78.7, 91.6)*	*73.2%*	*91.3%*	*6.0 (5.2–6.9)*	*85.0% (78.6, 89.7)*
Esat6 Ag	*33/50 (66.0%; 52.1, 77.6)*	*78/110 (70.9%; 61.8, 78.6)*	*50.8%*	*82.1%*	*2.3 (2.1–2.5)*	*69.4% (61.8, 76.0)*
38 kDa Ag	*47/50 (94.0%; 83.8, 97.9)*	*67/110 (60.9%; 51.6, 69.5)*	*55.2%*	*95.7%*	*2.4 (2.3–2.5)*	*71.2% (63.8, 77.7)*

Table 4. Sensitivity and specificity of 5 novel recombinant antigens by ELISA test [PTB = 111, EPTB = 29 and control = 110]. Pos: Positive, Neg: Negative, CI: Confidence interval, PPV: Positive predictive value, NPV: Negative predictive value, LRP: likelihood ratio for positive test, DA: Diagnostic accuracy. *Also see supplementary file 1.

To develope the serological tests, for the diagnosis of tuberculosis, researchers have used both crude as well recombinant antigens[8,9,12]. The performance of these tests has been evaluated using single as well as cocktail of multiple antigens[13–16]. The most commonly used and commercially available antigens have been the esat6 and *Mtb* 38 kDa recombinant antigens. However, this antigen showed sensitivity of only 47% and specificity of 94% in smear positive patients, in studies carried out earlier[13]. In present study also the performance of this antigen was far inferior to our newly described antigens. Another recombinant protein used earlier was malate synthase and MPT51. At laboratory level, the malate synthase showed the specificity of 98% and sensitivity of 73% in sputum smear positive patients[17,18]. The MPT51 protein showed poor sensitivity rates of 59% in HIV-negative TB patients and 58% in HIV-TB patients[19,20]. Its sensitivity and specificity rates reported by others were between 74 to 80% and 34 to 74% respectively[21,22]. Recently, Feng *et al.*[23] reported that a polyprotein comprising of 38 kDa and MPT64 are suitable for diagnosing active tuberculosis with sensitivity and specificity of 70.4% and 91.5% respectively. A recombinant TbF6 antigen was generated by fusion of four distinct antigens (38 kDa, CFP-10, MTB8, and MTB48) in a single protein. However, its sensitivity in sputum smear positive patients remained below 70% only[17,19,24,25]. These and several other antigens were used to develop commercial ELISA and rapid diagnostic test (RDT) kits, despite giving low sensitivity and specificity[3]. Prompted by concerns raised by academia in 2008, the WHO special programme for Research and Training in Tropical Diseases (TDR) performed an evaluation of 19 commercially available TB diagnostic kits[26]. The meta-analysis of these kit evaluation results showed the sensitivity of 1% to 60% and specificity of 53% to 99% under field conditions. Several other systematic reviews and original evaluation studies also showed similar findings[3–5,17,27].

Our antigens provided outstanding sensitivity and specificity, not reported earlier. This could most likely be due to the fact that in all previous studies recombinant antigens were prepared from the laboratory maintained H37Rv strain, while in our study we used a well characterized but fresh clinical isolate which developed *in-vivo* multidrug resistance[6,7]. We consider that this was a major turning point in the development of these antigens. It has been documented that expression levels as well as the characters of proteins expressed by the H37Ra, H37Rv, and clinical isolates are significantly different[28]. It is also important to mention that the diagnostic potential of the five recombinant proteins which we used in this study, has never been explored on such a scale for the diagnosis of tuberculosis[29].

For comparison we included esat6 and 38 kDa antigens as reference proteins. In fact while comparing the sensitivity and specificity of our antigens with these two reference antigens, we found that antigen rSS1, rSS4 and rSS5 are highly significantly superior to the esat6 and 38 kDa Ag or other previously published studies[30]. In comparison to reference antigen (esat6 and 38 kDa Ag), We noticed great difference in the specificity of rSS1, rSS2, rSS3, rSS4, rSS5 antigens, though only three antigens (rSS1, rSS4, rSS5) showed significant difference in the sensitivity. We found that in previous evaluation studies, several kits were marketed without evaluating the

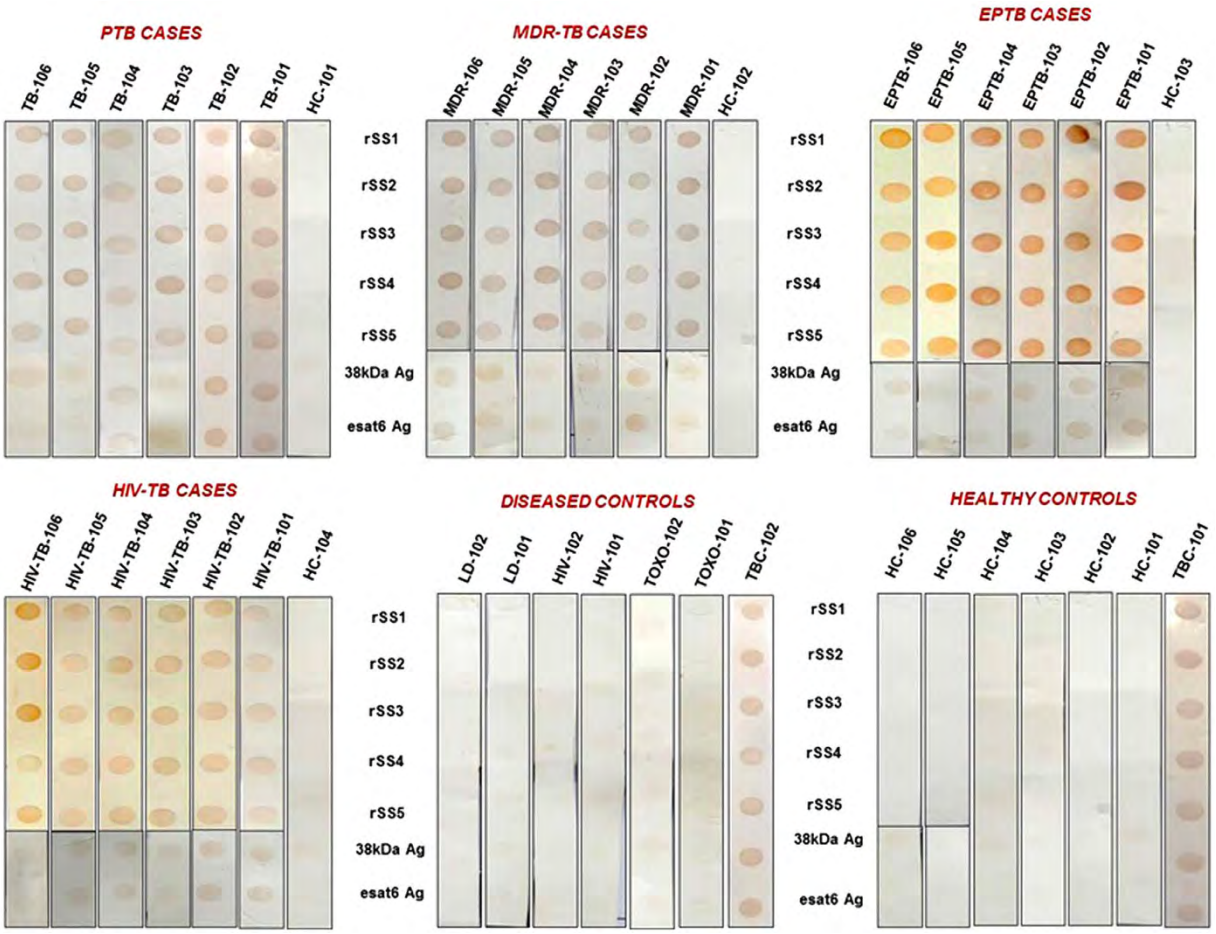

Figure 2. Dot-blot results showing the reactivity of purified recombinant proteins rSS1, rSS2, rSS3, rSS4, rSS5, Ag38 kDa and esat6 with serum samples from patients with extra-pulmonary tuberculosis (EPTB) pulmonary tuberculosis (PTB) HIV-TB, MDR-TB as well as from healthy controls (HC) and diseases controls (DC). Specific binding of the antigen to its antibody is indicated by a dark spot.

specificity of these antigens on disease controls (i.e. HIV, leishmaniasis, toxoplasmosis, cancer, and diabetes) in whom these commercial kits failed badly. However, we have included not only healthy controls but also various disease controls in this study. Most importantly, three recombinant antigens (rSS5, rSS1 and rSS4) showed very high utility in MDR-TB cases. This is an important finding of our study and we propose that these three antigens can be used not only for diagnosis of TB but can help in the ruling out the disease severity. We found higher sensitivity of all these antigen in dot-blot as compared to ELISA. The reason for this could be higher cut-off taken in ELISA while in dot-blot any visible dot was considered as positive. Likelihood ratio (LR) of positive test is also an important statistical method to better evaluate the diagnostic test. In our study, LR of positive test values was very high (ranging from 5.5–53.5). The areas under the ROC curve (AUC) showed an excellent diagnostic efficacy of our recombinant antigens (Table 4). The data also shows that our antigens do not cross react with *M. bovis* (BCG) vaccine. The antigen rSS4 is conserved hypothetical protein which stimulates T-cell response in the host. The rSS5 is a probable lipase/esterase LipN while rSS2 is an essential hypothetical protein[29]. The rSS3 is a phosphoglycerate kinase[6] involved in the second phase of glycolysis and to best of our knowledge, none of these antigen has been used for serodiagnosis of tuberculosis in India. Even though, the study showed very high utility of three antigen in the detection of active TB, but study had some limitations; all antigens were not tested with other disease control groups *i.e.* house hold contacts, latent TB, other pulmonary diseases such as asthma, bronchitis, pneumonia, allergies and also need to be evaluated in different geographical regions of world before commercialization.

Therefore, we can conclude that these results convince us that at least 3 of these five antigens can successfully be used for screening all the suspected cases of active tuberculosis, though not for confirming the diagnosis. We are in process of developing semi-quantitative rapid diagnostics tests (RDTs) in a way that these biomarkers can be used to detect and predicting tuberculosis, by measuring the high expression of these antigens. We strongly believe that rapid tests developed from these antigens can meet an urgent requirement of triage test for all forms of tuberculosis.

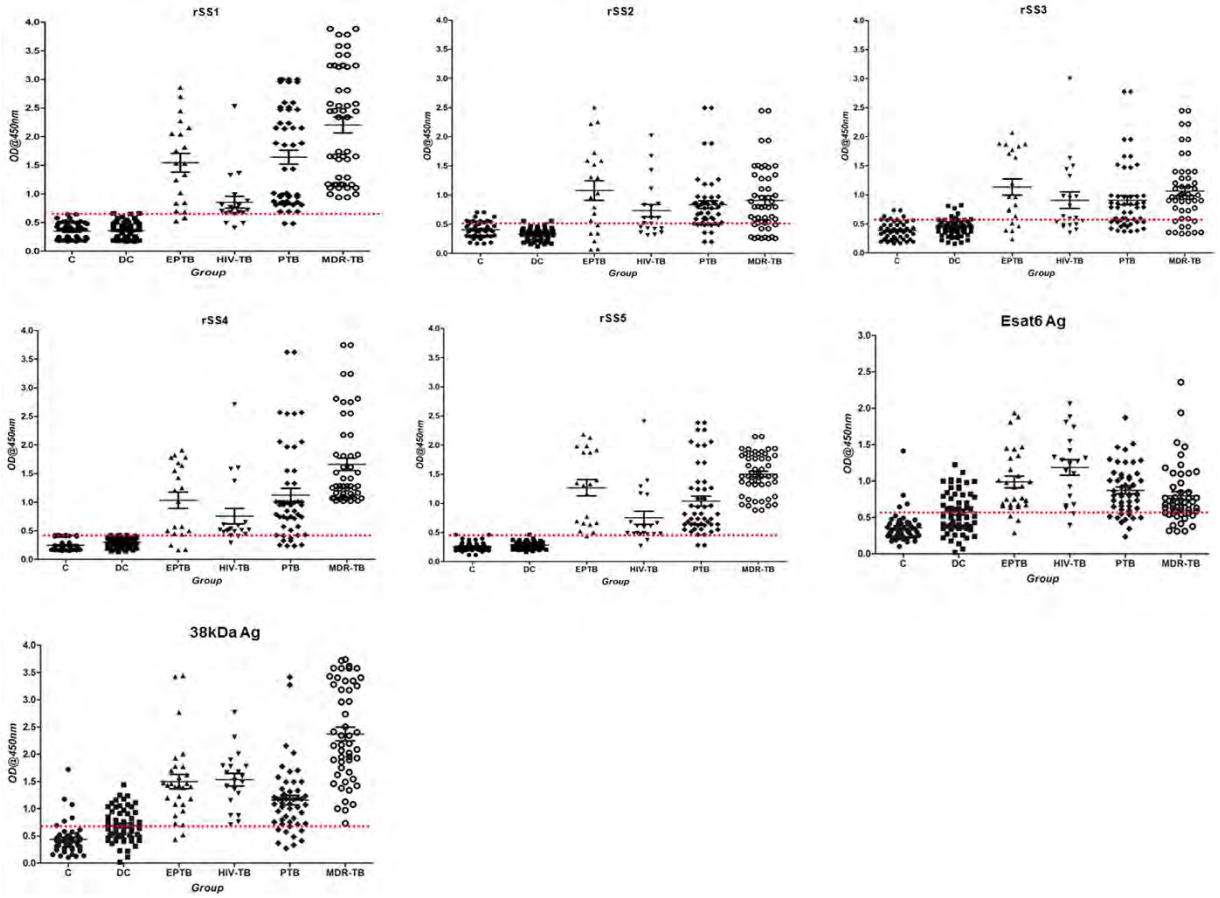

Figure 3. Scatter plots of ELISA results using our novel recombinant antigens and the reference antigens. Recombinant antigens assayed are rSS1, rSS2, rSS3, rSS4, rSS5, Esat-6 and 38 kDa Ag). The serum samples used were from of healthy controls (HC), diseased controls (DC), tuberculosis patient (PTB, EPTB, HIV-TB) and MDR-TB. The scatter plot indicates the antibody level per subject analysed. A dotted horizontal line is included to show the cut-off value for individual antigen.

Material and Methods

Cloning of the over-expressing genes. To express the desired antigens, the genomic DNA of *Mtb* clinical isolates (AIIMS/LM/SS/TB-1920/06) was used as the template for PCR reactions using gene specific primers (Table 1) as reported earlier[7]. The conditions for PCR were as follows: a 50 μl PCR mixture were comprised of 10 mM Tris-Cl (pH 8.2); 50 mM $MgCl_2$, deoxynucleoside triphosphate (200 μM each), 1.0 U of Taq DNA polymerase, 0.5 μl of each primer and 100 ng of genomic DNA. The reaction mixture was subjected to initial denaturation at 94 °C for 10 min followed by 30 amplification cycles of 94 °C for 1 min, annealing (rSS1 at 53 °C; rSS2 at 64 °C; rSS3 at 58 °C; rSS4 at 65 °C and rSS5 at 58 °C) for 45 sec, amplification at 72 °C for 45 sec and final extension at 72 °C for 10 min in a PTC-100 thermal cycler (MJ Research, USA).

Purified PCR product of each gene was cloned in to the pGEMT easy cloning vector and transformed in *E. coli* JM109 strain as per manufacturer instruction and plated out in appropriate antibiotic containing plates. Blue-white selection method was used for clones screening. Plasmids were extracted from randomly selected clones and subjected to restriction digestion by *HindIII* and *BamH*I restriction enzymes for clone confirmation. Gel purified inserts were sub-cloned into pQE-30 expression vector and transformed in *E. coli* M15 strain. All recombinant clones were selected in 100 μg/ml ampicillin and 50 μg/ml kanamycin containing Luria-Bertani (LB) agar plate. Plasmids were again extracted from all selected clones and presence of insert was confirmed by DNA sequencing.

Recombinant clones were cultured overnight at 37 °C in LB broth containing appropriate antibiotics. Overnight grown culture was inoculated into fresh LB medium containing antibiotics and incubated at 37 °C with shaking at 225 rpm until OD (600 nm) reached at 0.6. Thereafter, protein expression was induced by 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at the mid-exponential phase of growth. Cultures were grown for an additional 4 h with shaking at 37 °C for proteins expression. Bacterial pellets were harvested by centrifugation at 12,000 rpm for 5 min and stored at −80 °C until use. Pellets were re-suspended in lysis buffer (20 mM Tris (pH-8.0), 150 mM NaCl and 1 mM phenylmethylsulfonyl fluoride (PMSF) and disrupted by sonication followed by centrifugation at 12000 rpm for 30 min at 4 °C. Recombinant proteins were purified under native and denaturing conditions by using Ni^{2+} NTA metal-ion-affinity chromatography as per manufacturer's instructions (Qiagen, Germany). Purity level were analysed through sodium dodecyl sulphate polyacrylamide gel electrophoresis

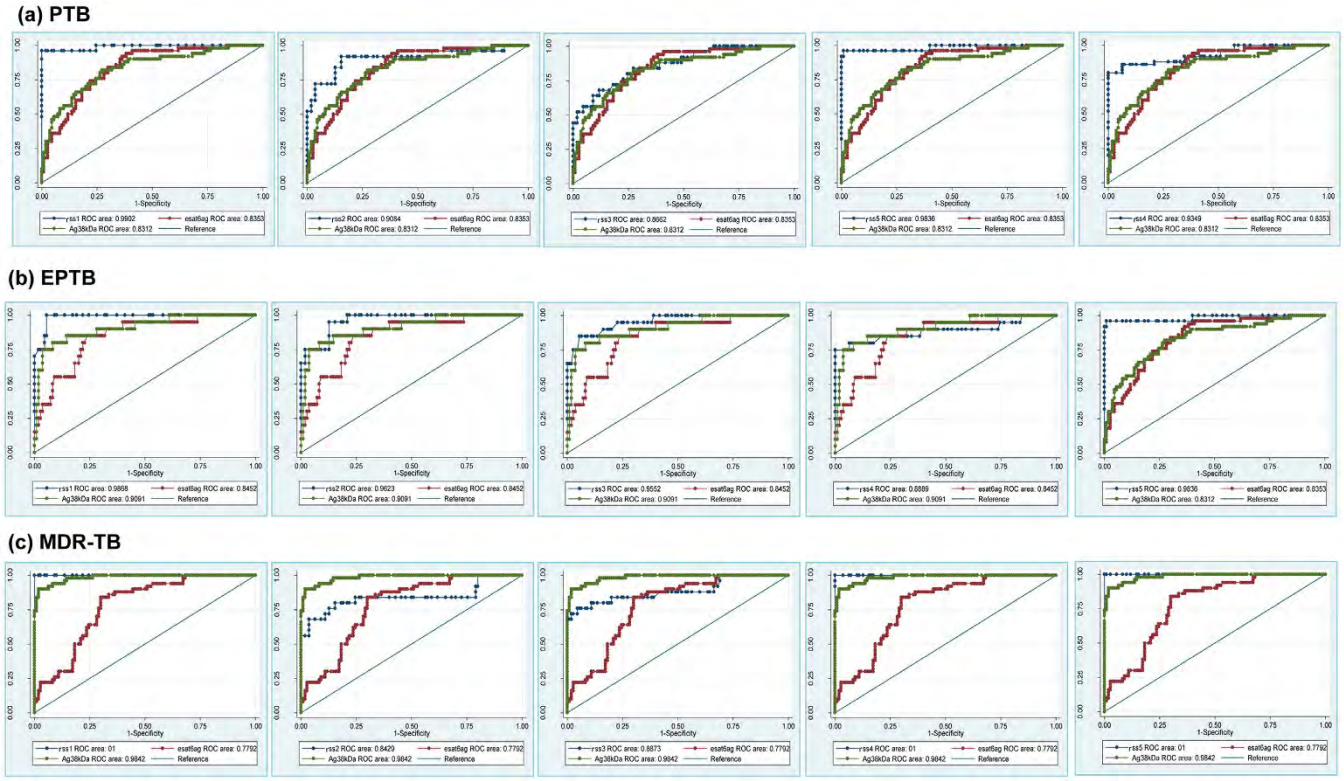

Figure 4. Receiver operative characteristic (ROC) curves of the antibody response against five *M. tuberculosis* recombinant antigens (rSS1, rSS2, rSS3, rSS4, rSS5) with two reference antigens (esat6 and Ag38 kDa) in (**a**) PTB, (**b**) EPTB and (**c**) MDR-TB patients and healthy controls.

(SDS-PAGE). Purified proteins were dialyzed against 20 mM Tris, pH 8.0, and concentrated using Amicon Ultra 10 kDa/3 kDa – cut off centrifugal filters (Millipore, India). The concentration of purified proteins was quantified by Bradford protein assay (Bio-Rad, USA).

Recombinant esat6 Ag and 38 kDa Ag proteins. The recombinant esat6 Ag and 38 kDa Ag were kind gifts from CIIDRET, University of Delhi South Campus, New Delhi. The preparation of esat6 is described previously[31,32]. The 38 kDa antigen was produced as N-terminal deca-histidine-tagged protein using T7 promoter-based expression system as described previously for hexa-histidine-tagged 38 kDa[33].

Study population and serum sample collection. The study was conducted from 2009–2015. Institutional ethics committee of the All India Institute of Medical Sciences (AIIMS), New Delhi approved the study (Ref No. T-9/31.7.2009). The participants were included after informed and written consent to participate in this study. A total of 250 subjects [TB patients (n = 140), DC (n = 60) and HC (n = 50)] were eligible for the evaluation of the diagnostic performance of five recombinant antigens. Of 140 culture confirmed TB cases, 111 were PTB PTB cases and 29 samples were from EPTB cases included 50 MDR-TB drug susceptibility (DST) confirmed cases (n = 46 PTB, n = 4 EPTB). The selection criteria for PTB cases was; similar to the diagnostic recommendation defined by Revised National Tuberculosis Control Program (RNTCP)[34]. In brief: was defined of a patient, presented with cough and fever >2 weeks, mysterious weight loss, fatigue, past history of patients/family members with TB or those diagnosed as having TB by the clinician and advised to receive full course of TB treatment. The medical history of the patients was obtained before collection of the blood sample and details of other investigations including chest x-ray. and other laboratory findings were recorded from the patient file. The clinically diagnosed case of TB, involving specimens from lung parenchyma or the tracheobronchial tree were classified as PTB, while clinical samples linking to organs other than the lungs, e.g. pleura, lymph nodes, abdomen, genitourinary tract, skin, joints and bones, meninges were classified as EPTB cases. The clinical isolates obtained from PTB and EPTB cases were subjected to first line drug susceptibility testing; those isolates showing resistance to isoniazid and rifampicin were classified as MDR-TB cases. Only three types of patients with HIV, Toxoplasmosis, Leishmaniasis (confirmed by commercial ICTs/ELISA based test) were included in the disease control group. The TB status in the participants of diseased group was confirmed by MGIT960 culture. Other patients with disease such as hepatitis, cancer, diabetes and other autoimmune diseases were excluded. The blood samples of all TB negative participants (healthy volunteers) were selected from the individuals with no sign/history of TB and who tested negative by tuberculin skin test (TST). The TST was performed by a trained phlebotomist. The 5 TU/0.1 mL tuberculin (Span Diagnostics, Surat, India) was administrated intradermally on the volar aspect of the forearm and read after 48–72 hours; an induration ≥10 mm was defined as positive. All

serum samples were stored in different aliquots and stored at −80 °C. Other patients details such as age, gender and vaccination and are listed in Table 2.

Western blot and dot-blot Assays. The purified proteins were resolved on 12% SDS-PAGE and transferred on to nitrocellulose membranes using semi-dry blotting apparatus (Bio-Rad, Hercules USA) following manufacturer instructions. For dot-blot, PBS was used as the spotting and dilution buffer for all purified recombinant proteins. A dot-blot was x spotted on a nitrocellulose membrane using a Bio-Dot 96-well manifold apparatus (Bio-Rad, Hercules, USA). The 100 μl of diluted proteins were loaded on each well of the assembled 96-well dot-blot apparatus with the final concentration of proteins 25 ng/well. The membrane was blocked with blocking buffer (5% skimmed milk in 1x Phosphate Buffer Saline (PBS)) for overnight at 4 °C. Membrane was washed 5 times with 1x PBST (1xPBS + 0.5% Tween20) and probed with culture confirmed PTB, EPTB or MDR-TB patients' sera (1:200 dilutions in 1xPBS containing 0.25% bovine serum albumin) for 2 hrs at 37 °C. Serum samples from healthy and disease controls were also tested in the same manner. After Incubation, membrane was again washed 5 times with 1x PBST and probed with anti-human IgG (whole molecule) HRP-conjugated antibody (Sigma, USA) as the secondary antibody (1:8000 dilutions) for 2 hrs at 37 °C. The membranes were developed by using 3,3′-Diaminobenzidine (DAB), (Sigma Aldrich, USA) as substrate through incubation at room temp for 30–40 second as substrate for 30–40 sec at RT.

Indirect Enzyme-Linked Immunosorbent Assay (ELISA). Checkerboard titration (CBT) method was used for the detection of optimum serum dilution and antigen concentration in ELISA. Ninety-six well flat bottom plates (BD, New Jersey, USA) were coated with purified antigens of 25 ng/well (rSS1 and rSS5) or 50 ng/well (rSS2, rSS3, rSS4, esat6 and 38 kDa Ag) in 100 μl/well 0.1 M bicarbonate buffer (pH 9.6) and incubated at 4 °C for overnight. Plates were washed 3 times with 1x PBST, blocked with blocking buffer (5% fat free skimmed milk in PBS) for two hour at 37 °C and followed by three washing with 1x PBST. Optimum serum dilutions (1:50) were added on each well and incubated for 2 h at 37 °C. The plates were again washed three times with PBST and then incubated for 2 h with anti-human IgG, whole antibody conjugated with horseradish peroxidase (1:15,000 dilutions) followed by washing with 1x PBST. Enzyme activity was assayed by incubation for 15 min at 37 °C with 100 μl of tetramethylbenzidine (TMB) per well. Reaction was stopped by using 50 μl of 2N sulfuric acid in each well and the OD was taken at 450 nm. All the serum samples were tested in duplicate wells and experiments were repeated three times to verify the reproducibility of results. All *in-vitro* methods were performed in accordance with the standard guidelines and institutional regulations and following the manufacturer's instructions, if a commercial equipment, device or kit was used.

Statistical analysis. Data were presented as means and standard deviation (mean ± 2 SD). The cut-off value of an ELISA was determined by using receiver operative characteristic (ROC) curve analysis. ROC curves were plotted using STATA SE.11.0 (Stata Corp LP, Texas, USA) software. ROC curve describes probability of accuracy of the test result at different cut-off values. The individual sample was scored positive, if the OD was above the Youden index (YI) value. Sensitivity was determined by dividing the number of positive cases by the total number of TB patients. Specificity was determined by dividing the number of negative controls by the total number of healthy controls.

References

1. WHO. *Global Tuberculosis Report*. WHO/HTM/TB/2015.22, (World Health Organization,Geneva, Switzerland, 2015).
2. Singh, S. & Katoch, V. Commercial serological tests for the diagnosis of active tuberculosis in India: Time for introspection. *Indian J. Med. Res.* **134,** 583 (2011).
3. Singh, S. *et al.* Poor Performance of Serological Tests in the Diagnosis of Pulmonary Tuberculosis: Evidence from a Contact Tracing Field Study. *PLoS ONE* **7,** e40213 (2012).
4. Steingart, K. R. *et al.* Commercial Serological Tests for the Diagnosis of Active Pulmonary and Extrapulmonary Tuberculosis: An Updated Systematic Review and Meta-Analysis. *PLoS Med* **8,** e1001062 (2011).
5. WHO STAG-TB. *WHO | Strategic and Technical Advisory Group for Tuberculosis (STAG-TB).* 1–25 (World Health Organization, 2011).
6. Singh, A. *et al.* Comparative proteomic analysis of sequential isolates of *Mycobacterium tuberculosis* from a patient pulmonary tuberculosis turning from drug sensitive to multidrug resistant. *Indian J Med Res* **141,** 27–45 (2015).
7. Singh, A., Gopinath, K., Singh, N. & Singh, S. Deciphering the sequential events during *in vivo* acquisition of drug resistance in *Mycobacterium tuberculosis*. *Int. J. Mycobacteriology* **3,** 36–40 (2014).
8. Raju, R. *et al.* Diagnostic role of the antibody response to the 38 kDa, 16 kDa proteins and lipoarabinomannan of *Mycobacterium tuberculosis*. *Indian J. Clin. Biochem. IJCB* **20,** 123–128 (2005).
9. Wang, X., Bao, L., Zhao, M., Zhang, H. & Long, Y. Expression of the fusion protein CFP10-ESAT6 of *Mycobacterium tuberculosis* and the study of its immunogenicity. *Sichuan Da Xue Xue Bao Yi Xue Ban* **37,** 353–356 (2006).
10. Perkins, M. D., Roscigno, G. & Zumla, A. Progress towards improved tuberculosis diagnostics for developing countries. *Lancet* **367,** 942–943 (2006).
11. Malhotra, V. L., Lakshmy, A. & Nagarkar, M. Clinical utility of EIA in serodiagnosis of extra-pulmonary tuberculosis using 38 kDa mycobacterial antigen. *J. Commun. Dis.* **43,** 281–284 (2011).
12. Zhang, S.-L. *et al.* Development and evaluation of a novel multiple-antigen ELISA for serodiagnosis of tuberculosis. *Tuberc. Edinb. Scotl.* **89,** 278–284 (2009).
13. Silva, V. M. C., Kanaujia, G., Gennaro, M. L. & Menzies, D. Factors associated with humoral response to ESAT-6, 38 kDa and 14 kDa in patients with a spectrum of tuberculosis. *Int. J. Tuberc. Lung Dis.* **7,** 478–484 (2003).
14. Abebe, F., Holm-Hansen, C., Wiker, H. G. & Bjune, G. Progress in serodiagnosis of *Mycobacterium tuberculosis* infection. *Scand. J. Immunol.* **66,** 176–191 (2007).
15. Bukhary, Z. A. Evaluation of anti-A60 antigen IgG enzyme-linked immunosorbent assay for serodiagnosis of pulmonary tuberculosis. *Ann. Thorac. Med.* **2,** 47–51 (2007).
16. Hoff, S. T. *et al.* Evaluation of *Mycobacterium tuberculosis* specific antibody responses in populations with different levels of exposure from Tanzania, Ethiopia, Brazil, and Denmark. *Clin. Infect. Dis.* **45,** 575–582 (2007).

17. Houghton, R. L. *et al.* Use of multiepitope polyproteins in serodiagnosis of active tuberculosis. *Clin. Diagn. Lab. Immunol.* **9,** 883–891 (2002).
18. Singh, K. K. *et al.* Combined use of serum and urinary antibody for diagnosis of tuberculosis. *J. Infect. Dis.* **188,** 371–377 (2003).
19. Ramalingam, B., Uma Devi, K. & Raja, A. Isotype-specific anti-38 and 27 kDa (MPT 51) response in pulmonary tuberculosis with human immunodeficiency virus coinfection. *Scand J Infect Dis* **35,** 234–239 (2003).
20. Bethunaickan, R., Baulard, A., Locht, C. & Raja, A. Antibody response in pulmonary tuberculosis against recombinant 27 kDa (MPT51, Rv3803c) protein of *Mycobacterium tuberculosis*. *Scand J Infect Dis* **39,** 867–874 (2007).
21. Wu, X. *et al.* Humoral Immune Responses against the *Mycobacterium tuberculosis* 38-Kilodalton, MTB48, and CFP-10/ESAT-6 Antigens in Tuberculosis. *Clin. Vaccine Immunol.* **17,** 372–375 (2010).
22. Cheng, Z. *et al.* Evaluation of a novel fusion protein antigen for rapid serodiagnosis of tuberculosis. *J. Clin. Lab. Anal.* **25,** 344–349 (2011).
23. Feng, X. *et al.* Enhanced serodiagnostic utility of novel *Mycobacterium tuberculosis* polyproteins. *J. Infect.* **66,** 366–375 (2013).
24. Murthy, M., Parasa, R., Deenadayalan, A., Sharma, P. & Raja, A. Evaluation of the diagnostic potential of region of deletion-1-encoded antigen culture filtrate protein-10 in pulmonary tuberculosis. *Diagn Microbiol Infect Dis* **59,** 295–302 (2007).
25. Ramalingam, B., Baulard, A. R., Locht, C., Narayanan, P. R. & Raja, A. Cloning, expression, and purification of the 27 kDa (MPT51, Rv3803c) protein of *Mycobacterium tuberculosis*. *Protein Expr. Purif.* **36,** 53–60 (2004).
26. WHO. *Laboratory-based evaluation of 19 commercially available rapid diagnostic tests for tuberculosis.* (World Health Organization on behalf of the Special Programme for Research and Training in Tropical Disease Switzerland (2008).
27. WHO. *Policy statement: commercial serodiagnostic tests for diagnosis of tuberculosis.* 1–26 (World Health Organization, 2011).
28. Jhingan, G. D. *et al.* Comparative proteomic analyses of avirulent, virulent, and clinical strains of *Mycobacterium tuberculosis* identify strain-specific patterns. *J. Biol. Chem.* **291,** 14257–14273 (2016).
29. Eweda, G., Suzuki, D., Nagata, T., Tsujimura, K. & Koide, Y. Identification of murine T-cell epitopes on low-molecular-mass secretory proteins (CFP11, CFP17, and TB18.5) of *Mycobacterium tuberculosis*. *Vaccine* **28,** 4616–4625 (2010).
30. Zhang, C. *et al. Mycobacterium tuberculosis* secreted proteins as potential biomarkers for the diagnosis of active tuberculosis and latent tuberculosis infection. *J. Clin. Lab. Anal.*. doi: 10.1002/jcla.21782 (2014).
31. Kulshrestha, A. *et al.* Expression and purification of recombinant antigens of *Mycobacterium tuberculosis* for application in serodiagnosis. *Protein Expr. Purif.* **44,** 75–85 (2005).
32. Singh, V., Kaur, C., Chaudhary, V. K., Rao, K. V. S. & Chatterjee, S. *M. tuberculosis* secretory protein ESAT-6 induces metabolic flux perturbations to drive foamy macrophage differentiation. *Sci. Rep.* **5,** 12906 (2015).
33. Chaudhary, V. K. *et al.* Expression and purification of recombinant 38-kDa and Mtb81 antigens of *Mycobacterium tuberculosis* for application in serodiagnosis. *Protein Expr. Purif.* **40,** 169–176 (2005).
34. Chauhan, L. S. & Arora, V. K. Central TB Division, Directorate General of Health Services, Ministry of Health and Family Welfare & Indian Academy of Pediatrics. Management of pediatric tuberculosis under the Revised National Tuberculosis Control Program (RNTCP). *Indian Pediatr.* **41,** 901–905 (2004).

Acknowledgements

The research was supported by grant from Department of Biotechnology, Government of India, New Delhi (BT/PR7880/MED/14/1161/2006) and Indian Council of Medical Research, New Delhi (5/8/5/11/2010-ECD-I). A.S. was supported by a Senior Research fellowship from the Indian Council Medical Research (ICMR), Government of India. The authors are thankful to Prof. Vijay Kumar Chaudhary, Centre for Innovation in Infectious Disease Research, Education and Training (CIIDRET), University of Delhi South Campus, Delhi for providing purified recombinant 38 kDa and esat6 proteins free of cost. Authors acknowledge the help extended by Dr. M. Kalaivani, Scientist-III, Department of Biostatistics, All India Institute of Medical Sciences, New Delhi in statistical analysis of the data. Authors also wish to thank Saumya Srivastava, Jitendra Singh, Nand Kishore, Shruti Kahol, Vinod Kumar, Kriti Upadhyay, Ayush Chaudhary, Manisha Jain, and Veena Balooni (all from Division of Clinical Microbiology & Molecular Medicine, All India Institute of Medical Sciences, New Delhi), for their help in the study.

Author Contributions

A.S. and A.K.G. carried out the experiments, participated in the data analysis. A.S., K.G. and S.S. conceived and designed the study, A.S., A.K.G. and S.S. interpreted the experiment data and drafted the manuscript. All authors read and approved the final manuscript. S.S. and P.S. contributed reagents/materials/analysis tools.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Singh, A. *et al.* Evaluation of 5 Novel protein biomarkers for the rapid diagnosis of pulmonary and extra-pulmonary tuberculosis: preliminary results. *Sci. Rep.* **7**, 44121; doi: 10.1038/srep44121 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2017

Indian J Med Res 141, January 2015, pp 27-45

Comparative proteomic analysis of sequential isolates of *Mycobacterium tuberculosis* from a patient with pulmonary tuberculosis turning from drug sensitive to multidrug resistant

Amit Singh[1,¶], Krishnamoorthy Gopinath[1,¶,†], Prashant Sharma[2], Deepa Bisht[2], Pawan Sharma[3,††], Niti Singh[4] & Sarman Singh[1#]

[1]Division of Clinical Microbiology & Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi, [2]National JALMA Institute for Leprosy & other Mycobacterial Diseases (ICMR), Agra, [3]Immunology Group, International Centre for Genetic Engineering & Biotechnology, New Delhi & [4]Department of Microbiology, National Institute of Tuberculosis & Respiratory Diseases, New Delhi, India

Received September 10, 2013

***Background & objectives*: Tuberculosis is a major health problem in India, and the emergence of multidrug resistant (MDR) and extensively drug resistant (XDR) strains of *Mycobacterium tuberculosis* (*Mtb*) has further complicated the situation. Though several studies characterizing drug sensitive and drug resistant strains are available in literature, almost all studies are done on unrelated strains. Therefore, the objective of this study was to compare the proteomic data of four sequential isolates of *Mtb* from a single patient who developed MDR-TB during the course of anti-tuberculosis therapy (ATT).**

***Methods*: In this study, using two-dimensional (2D) gel electrophoresis and MALDI-TOF mass spectrometry, we compared and analyzed the cell lysate proteins of *Mtb* sequential clinical isolates from a patient undergoing anti-TB treatment. The mRNA expression levels of selected identified proteins were determined by quantitative real-time polymerase chain reaction (qRT-PCR).**

***Results*: The genotypes of all four isolates remained homologous, indicating no re-infection. The initial isolate (before treatment) was sensitive to all first-line drugs, but the consecutive isolates were found to be resistant to isoniazid (INH) and rifampicin (RIF) and developed mutations in the *katG*, *inhA* and *rpoB*. The intensities of 27 protein spots were found to be consistently overexpressed in INH and RIF resistant isolates. The most prominent and overexpressed proteins found during the development of drug resistance were GarA (Rv1827), wag31 (Rv2145c), Rv1437 and Rv2970c.**

***Interpretation & conclusions*: This preliminary proteomic study provides an insight about the proteins that are upregulated during drug resistance development. These upregulated proteins, identified here, could prove useful as immunodiagnostic and possibly drug resistant markers in future. However, more studies are required to confirm these findings.**

Key words 2D gel electrophoresis - tuberculosis - MDR-TB - MALDI-TOF - proteomics

¶Authors contributed equally

Present address: †Institute of Infectious Diseases & Molecular Medicine, University of Cape Town, South Africa

††North-East Region-BPMC, Department of Biotechnology, Govt. of India, New Delhi

Tuberculosis (TB) is a global emergency with an estimated nine million new cases and more than 1.5 million deaths occurring annually[1]. The situation has worsened after AIDS epidemic and with the emergence of multidrug resistant (MDR) forms of the causative agent *Mycobacterium tuberculosis* (*Mtb*). These drug resistant strains are more infectious by virtue of their high transmissibility in the population. Therefore, identification of the reliable diagnostic, prognostic and drug resistance markers is an urgent research priority. Various *in vitro* and *in vivo* studies have identified chromosomal mutations as determinants of drug resistance[2-4]. For example, mutation (s) in *rpoB* allele confers rifampicin (RIF) resistance (RIF^r) in 90-95 per cent isolates[2], while isoniazid (INH)-resistance (INH^r) is attributed to mutation (s) in one or more alleles *viz., katG, inhA, ahpC* and *ndh*. However, in about 20 per cent of INH^r isolates, none of these known mutations are found, suggesting the possibility of unknown mutations or mechanisms[2,3]. On the other hand, the genetic mutations may not necessarily correlate with phenotypic resistance; further suggesting that other factors such as drug impermeability, drug-efflux pumps, formation of survivable "persister cells" under drug pressure and several other host factors could be involved in the outcome of treatment[2].

Hence, for unraveling the mechanism(s) of drug resistance, understanding the mode of action of anti-TB drugs is very crucial. Many studies have elucidated the mode of action of various anti-TB drugs using genetic analysis, mRNA expression and DNA microarray analysis[4,5]. Several groups have also explored the proteome of *Mtb* and provided comprehensive details about the subcellular localization and confirmed the genomic annotation[6-9]. In these studies two-dimensional (2D) gel electrophoresis followed by mass spectrometry (MS) identification of the differentially regulated proteins substantially helped in identifying the complex pathways and their regulatory enzymes. These studies also elucidated modes of action of various drugs and discovered new antigens that could be potential candidates for developing vaccines and diagnostics[6,7,9,10]. However, only a few studies are available which show differential expression of specific proteins in the drug resistant but not in drug susceptible cells[7–9,11]. Further, in all these studies, either the non-pathogenic mycobacteria or laboratory collections of drug sensitive and drug resistant strains of *Mtb* from different patients have been used. In the present study, protein profile of sequentially collected four clinical isolates of *Mtb* was analyzed using 2D gel electrophoresis and the differentially expressed proteins were identified by MALDI-TOF-MS analysis. All isolates were from the same patient, who developed MDR-TB during the course of chemotherapy.

Material & Methods

The study was conducted between January 2006 and June 2010 at the TB Laboratory, Division of Clinical Microbiology and Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi, India. This study was approved by the Institutional Ethics Committee of AIIMS and written informed consent was obtained from the patient. The patient was being treated at the designated microscopy and DOTS (directly observed treatment-short course) centres of Shahpurjat, New Delhi. This patient (22 yr old male) was diagnosed as having pulmonary TB on the basis of clinical and radiological findings and sputum smear microscopy. He was prescribed with anti-TB treatment (ATT) under the DOTS programme. The thrice a week treatment regimen comprised isoniazid, rifampicin, pyrazinamide (PZA) and ethambutol (EMB) (category I treatment) in intensive phase for two months followed by four month treatment with two (isoniazid and rifampicin) drugs regimen. Pre-treatment sputum specimen was used for isolation of *Mycobacterium* sp. by BACTEC MGIT-960 (Becton Dickinson, Sparks, MD, USA), which was positive. The isolate was identified as *Mtb* by conventional phenotypic and in-house PCR method[12]. This culture was labelled as isolate A, and was subjected to 16sRNA gene sequencing. The patient though took full six months course of treatment but became irregular in taking drugs after initial improvement in his clinical symptoms. After three months of cessation of treatment (6+3= 9 month[13], his condition again deteriorated and his sputum culture was again positive for *Mtb*. We labelled this second culture as isolate B. He was re-treated with isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin (SM) (category II regimen). Within two months his clinical condition improved but he again defaulted. After an asymptomatic period of about four months his symptoms reappeared. His sputum was again culture positive and this culture was labelled as isolate C. The patient was again prescribed with the same treatment for 12 months after counselling but he stopped treatment after six months. His condition further deteriorated and he died of multisystem failure. The fourth sample was received just before his death

and the isolate from this sputum sample was labelled as isolate D.

All the four clinical isolates (A, B, C & D) were identified as *Mtb* using standard protocols[12,13]. The anti-mycobacterial drug susceptibility testing was performed on all the isolates by both BACTEC™ MGIT-960 (Becton Dickinson, Sparks, MD, USA) and proportional method using Middlebrook 7H10 (Difco, USA) agar plates containing first-line anti-TB drugs (SM 2.0 µg/ml, INH 0.2 µg/ml, RIF 1.0 µg/ml, EMB 6.0 µg/ml)[13,14]. All four isolates were also genotyped by spoligotyping and identified using SITVIT-WEB database[15]. The *rpoB, inhA* and *katG* gene targets were sequenced using the primers as described elsewhere[13].

Preparation of mycobacterial whole cell lysate: All *Mtb* isolates were grown without shaking in Middlebrook 7H9 medium supplemented with 0.2 per cent (v/v) glycerol, 10 per cent oleic acid, albumin-dextrose and catalase (OADC, Difco, USA) at 37°C for two weeks. Whole cell lysate was prepared according to protocol of Sharma *et al*[11]. Cells were washed three times with normal saline and then suspended in sonication buffer [50 mM tris-HCl containing 10 mM $MgCl_2$, 0.1% sodium azide, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1mM ethylene glycol tetra acetic acid (EGTA); *p*H 7.4] at a concentration of 1g wet cell mass per 5ml, and then broken by intermittent sonication for 15 min at 4°C using sonicator (Sonics & Materials Inc, USA). The homogenate was centrifuged at 12,000×g for 20 min at 4°C. The pellets were discarded and supernatant was stored at -70°C until further use.

Protein precipitation with sodium dodecyl sulphate (SDS)-trichloroacetic acid (TCA)-acetone: The cell lysates were treated with 1 per cent SDS and then subjected to TCA-acetone precipitation procedure[9]. The protein pellet was suspended in appropriate volume of two-dimensional rehydration buffer (Bio-Rad, USA), and the protein concentration was estimated using the Bradford method[16].

Two-dimensional gel electrophoresis: Isoelectric focusing (IEF) was done using the in-gel rehydration method (Bio-Rad, USA).

2D gels were analysed using PDQuest Advanced software (version 8.0) (Bio-Rad, USA). After acquisition, the images were analyzed using step-wise spot detection and spot matching followed by differential expression analysis. The quantity of each spot was normalized by total valid spot intensity. The expression differences for all four mycobacterial isolates were compared using the same software. Images for sensitive and resistant isolates were manually checked for artifactual spots, merged spots and missed spots, and spots with more isolate-specific variability were omitted in the downstream processing. Equal amount of protein was loaded in all gels and experiments were repeated three times with three independent biological replicates.

In-gel digestion of protein spots with trypsin: Protein spots of interest were excised from the coomassie brilliant blue R250 stained 2D gels using spot picker Investigator ProPic (Genomic Solutions Ltd., Huntingdon, UK). Digestion of proteins and spotting of peptides on matrix assisted LASER desorption/ ionization-time of flight (MALDI-TOF) target plate was carried out using protein digester investigator ProPrep (Genomic Solutions, Huntingdon, UK).

For protein digestion, method of Shevchenko *et al*[17] was followed with slight modifications. In brief, the gel plugs were de-stained and dehydrated by washing three times (10 min) with 25 mM NH_4HCO_3-50 per cent acetonitrile (ACN) (1:1v/v) solution and treated with freshly prepared 10 mM dithiothreitol (DTT) in 50 mM NH_4HCO_3 for 45 min at 56°C. After incubation, DTT was replaced with freshly prepared 55 mM iodoacetamide for 30 min and then dehydrated with 100 per cent ACN. The dried gel pieces were incubated for 12 h at 37 °C with 25 mM NH_4HCO_3 containing 0.02 µg/µl of mass spectrometry grade trypsin (Promega, USA). The resulting peptides were extracted twice from the gel pieces, using peptide extraction buffer [1:1v/v mixture of 70% ACN and 0.1% trifluoroacetic acid (TFA)].

Mass spectrometric analysis: Mass spectrometry (MS) was carried out as described earlier[9]. The digested samples were desalted and concentrated on C-18 ZipTips (Millipore, USA) using the manufacturer's protocol before mass spectrometric analysis. ZipTips were eluted on MTP 384 target plate with 2 µl of α-cyano-4-hydroxycinnamic acid (HCCA) (Sigma-Aldrich, USA) saturated solution dissolved in 50 per cent ACN and 0.2 per cent TFA. Mass spectra of digested proteins were acquired using Autoflex II TOF/TOF 50 (Bruker GmbH, Leipzig, Germany) in positive reflectron mode. AnchorChip target plate was placed in sample inlet of the instrument, controlled by flexControl 2.4 software (Bruker, Germany). The instrument was equipped with a 337 nm nitrogen LASER, delayed extraction

electronics, and a 50Hz digitizer and percentage of LASER energy was maintained at 30-40 per cent. The pulse energy was 105 µJ and pulse duration was 1.3 nano sec. Final mass spectra were produced by averaging 1500-2500 LASER shots taken at different positions within each spot. The spectra were acquired in positive reflection mode in the mass range of 500-3000 m/z. Calibration was performed using peptide calibration standard II (Bruker, Germany). The proteolytic masses obtained, were processed through FlexAnalysis v.2.4 programme (Bruker, Germany) for peak detection. Initially, the spectra acquired were processed for baseline subtraction with 80 per cent baseline flatness followed by smoothening with threshold of signal-to-noise ratio (S/N) >5. The contaminant m/z peaks originating from human keratin, trypsin auto-digestion and matrix were removed from the spectra to generate the peptide mass list for the database search. Finally, the proteolytic masses obtained were evaluated using MASCOT, a peptide mass fingerprinting (PMF) tool (Matrix Sciences, UK). Peak detection in MALDI spectra and peak lists were submitted to the UniProtKB/Swiss-Prot database using the MASCOT search engine (*http://www.matrixscience.com*) to identify the proteins from the annotated *Mtb* chromosome (strain H37Rv, EMBL/GenBank/DDBJ entry AL123456). Peptide mass tolerance was set in range of 50-100 ppm, with carbamidomethyl-cystein set as fixed modification, oxidation of methionine as variable modification and only 0 or 1 missed cleavage site was allowed. Further, matched precursor ions of identified proteins were selected for subsequent fragmentation using post source decay (PSD) for MS/MS. Lift_ATT method was performed in flex control software; parent peak mass spectrum was acquired by hitting LASER for 400-550 shots followed by acquisition of fragments of selected precursor ions for the same number of shots. Both parent and fragment spectra were pooled to generate MS/MS spectrum of a particular peptide. MS/MS spectrum was submitted to database using MASCOT wizard (Matrix Sciences, UK). The same parameters were used for MS/MS search in addition to the fragment mass tolerance from 0.5 to 2.0 Da.

Glycogen estimation: Logarithmic and stationary phase growth of *Mtb* sensitive (isolate A) and MDR (isolates B,C,D) isolates were collected by centrifugation (3 min at 5000 x *g* and at 4°C), and the pellet (15-20 mg wet weight) was re-suspended in 0.25M Na_2CO_3 and incubated at 95°C for 4 h. The glycogen content was estimated by following the procedure of Schulze *et al*[18].

Isolation of Mtb total RNA and real-time quantitative PCR (qRT-PCR): *Mtb* H37Rv was grown in Middlebrook 7H9 broth containing 10 per cent OADC, and was treated with INH (0.1µg/ml), RIF (1.0µg/ml), EMB (5µg/ml) and INH (0.1µg/ml) +RIF (1.0µg/ml). Total RNA was isolated using a TRI reagent (Sigma, USA) following manufacturer's instructions. To analyze mRNA expression, cDNA was synthesized from 1 µg of total RNA by using Superscript III (Invitrogen Life Technologies, USA) and random primers (Invitrogen, Life Technologies, USA), followed by amplification of the gene(s) by gene-specific primers, using master mix SYBR green (Applied Biological Materials Inc., Canada). The expression is represented in fold increase. 16sRNA was used as internal control for mRNA expression analysis of *ahpC*, *Rv1827*, *pknA*, *pknB*, *pknG* and *wag31*.

Each reaction was repeated thrice with three independent RNA samples in a smart cycler Cepheid machine (Cepheid, USA). RT-PCR conditions were as follows: an initial denaturation step of 10 min, followed by 40 amplification cycles of 30 sec at 95°C, 30 sec at 60°C and 30 sec at 72°C. Melting curve analysis was carried out to confirm the specificity of the amplified product. After baseline corrections and determination of threshold settings, calculation and statistical analyses were carried out using the $2^{-\Delta\Delta C}{}_{T}$ Method[19]. The results are shown as fold increase in expression profile.

Results

Drug resistance pattern and mutations in katG, inhA and rpoB genes of the isolates: The four sequential culture isolates were identified as *Mtb* by conventional phenotypic and in-house PCR method. Genotyping was done by spoligotyping and the results confirmed that all the isolates belonged to the Central Asian Strain Delhi (CAS1_Delhi, ST26) genotype. The drug susceptibility test results showed that the initial isolate (isolate A) was sensitive to all the four first-line drugs (SM, INH, RIF and EMB), but the consecutive isolates (isolates B, C and D) became resistant to three drugs; INH, RIF and EMB. The minimum inhibitory concentration (MIC) of isolate B increased as compared to isolate A against INH, RIF and EMB. However, it was still sensitive to kanamycin. The isolates C and D became resistant not only to INH, RIF and EMB but also to kanamycin. The sequencing of the *rpoB* (RIF^r), *katG* and *inhA* (INH^r) regions revealed mutated alleles associated with resistance to the respective drugs[13]. Morphologically the resistant isolates were stunted, thicker and cocco-bacillary in shape.

Differentially expressed proteins in drug sensitive and resistant isolates: The cell lysate proteins of four *Mtb* isolates were analyzed by 2D gel electrophoresis, which showed 430 protein spots in isolates A and 495, 556 and 395 spots in isolates B, C and D, respectively (Figs 1 & 2). Quantitative analysis of 2D gel spots was carried out using PDQuest software which revealed 27 spots upregulated in MDR isolates (Table I). The spots showing more than 2-fold upregulation were further identified by MALDI-TOF/TOF MS (Table II). To rule out possibility of any artifact, proteins showing equal intensity were taken as internal control (represented as square in Fig. 1). Upregulated proteins were functionally classified according to TubercuList web server which showed that most of the identified proteins belonged to the functional group 0, 1, 2, 3, 5, 7 and 9; corresponding to virulence, detoxification and adaptation (18.5%), lipid metabolism (11.11%), information pathway (14.81%), cell wall and cell process (3.7%), insertion sequence and phages (18.51%), and intermediary metabolism and information (29.62%); respectively (Table III). The magnified regions of upregulated proteins are shown in Fig. 3. Of the 27 upregulated proteins, eight were hypothetical protein (Rv2004c), probable glutamyl-tRNA (GLN) amidotransferase A gatA (Rv3011c), possible phosphoserine aminotransferase SerC (Rv0884c), probable lipase/esterase LipN (Rv2970c), probable phosphoglycerate kinase Pgk (Rv1437), conserved hypothetical protein with FHA domain, GarA (Rv1827), bacterioferritin (Rv1876) and conserved hypothetical protein (Rv0543) and were not found in 2D-PAGE database system accessible at *http://www.mpiib-berlin.mpg.de/2D-PAGE*, whereas three proteins probable iron-regulated aconitate hydratase Acn (Rv1475c), probable chaperone protein DnaK (Rv0350) and 60 kDa chaperonin 2 groEL2 (Rv0440), were found in two spots. Six proteins (gatA, serC, fbd, garA, Rv2204c and Rv0543c) in isolate B, 10 proteins (Rv685c, Rv3457c, Rv1479, Rv2970c, Rv1437, qor, and two spots each of Rv1475c and groEL2 family) in isolate C, three proteins (fadB, fabG4 and rrf) in isolate D, three proteins (Rv3075c, Rv1436 and GroES) in isolates C as well as in D were found upregulated. Only five proteins were consistently upregulated in all the three resistant isolates and these were identified as, chaperonin protein dnaK HSP70 (spots 4 and 5), hypothetical protein (Rv2004, spot 8), antigen 84 (wag31, spot 19) and bfrA (spot 24) (Fig. 1).

Among the identified proteins, we were more interested in studying the possible role of Rv1827 (GarA) and wag31 in drug resistance, since these proteins have been identified as physiological substrates for protein kinases G (pknG). Our result revealed that GarA and wag31 were upregulated in the drug resistant isolates. We analyzed the mRNA expression of *Rv1827* and its cognate protein kinases, *pknG, pknB, pknA* and *wag31*.

Drug induced changes in mRNA expression of protein kinases: To verify our protein expression observations, we studied the mRNA expression profile to see the effect of the four drugs on the standard strain of *Mtb* (H37Rv) which was sensitive to all anti-TB drugs. For this, the mRNA from H37Rv strain was isolated before and after exposing it to INH (0.1µg/ml), EMB (5.0 µg/ml), RIF (1.0 µg/ml) and INH+RIF (0.1 + 1.0µg/ml) for 6 h. Consistent with the proteomic data, seen in clinical isolates, *Rv1827* expression was upregulated in all the tested conditions. As expected, the upregulation was 6.82 fold when the *Mtb* standard strain (H37Rv) was exposed to INH and RIF together, but other tested genes had relatively diminished expression. While combining the EMB, the expression of *wag31* was higher and *pknA* and *pknG* expressions were highest (Fig. 4).

Glycogen storage: GarA, which is a glycogen regulatory protein, was found upregulated in our MDR isolates. It was found that as compared to sensitive isolates the glycogen accumulation in MDR isolates was higher. Consistent with GarA protein levels, the glycogen accumulation measured after seven days was 1.8, 2.0 and 2.1 folds higher in isolates B, C and D, respectively, as compared to sensitive isolate A. Interestingly, after 15 days the glycogen storage remained almost unchanged (Fig. 5).

Discussion

Emergence of drug resistance in *Mtb* has become a major concern for TB control programme managers and treating physicians. Though advances in genome sequencing methods have provided better opportunities to our understanding about functional genomics and proteomics of the *Mtb*, the knowledge about mechanism of drug resistance still remains limited only to the association of genetic polymorphism. Most often, the data from proteomic studies are used to understand host-pathogen interaction, virulence, drug resistance and drug tolerance[8,9,19-23]. Such studies have provided a comprehensive list of *Mtb* proteins that are found differentially regulated in laboratory maintained standard H37Rv strain exposed to drug pressure.

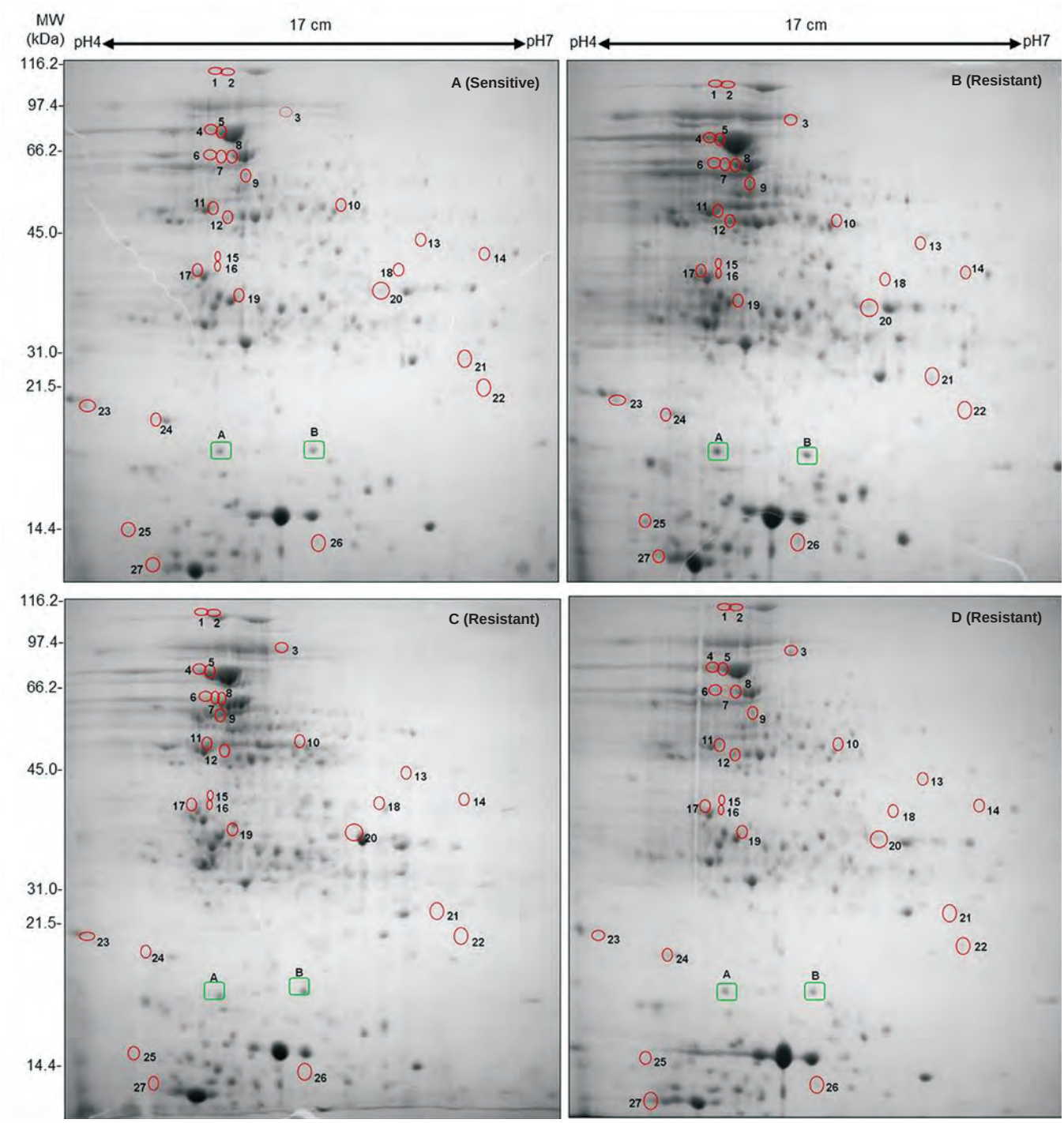

Fig. 1. 2D gel profiles of whole cell lysate proteins of *Mtb* clinical isolates collected sequentially from a single patient. The upregulated proteins are highlighted by circles. **(A)** First *Mtb* isolate (before treatment) sensitive to all 4 drugs, **(B)** Second isolate (during treatment) acquires MDR, **(C)** Third isolate (after 15 months of treatment) acquires drug resistance to yet another drug kanamycin and **(D)** Fourth MDR isolate after 27 months. Two Proteins, A (Rv1080c) and B (Rv2140c) marked in green rectangles, were selected for observing expression variation in all four samples and showed similar level of expression in all gels.

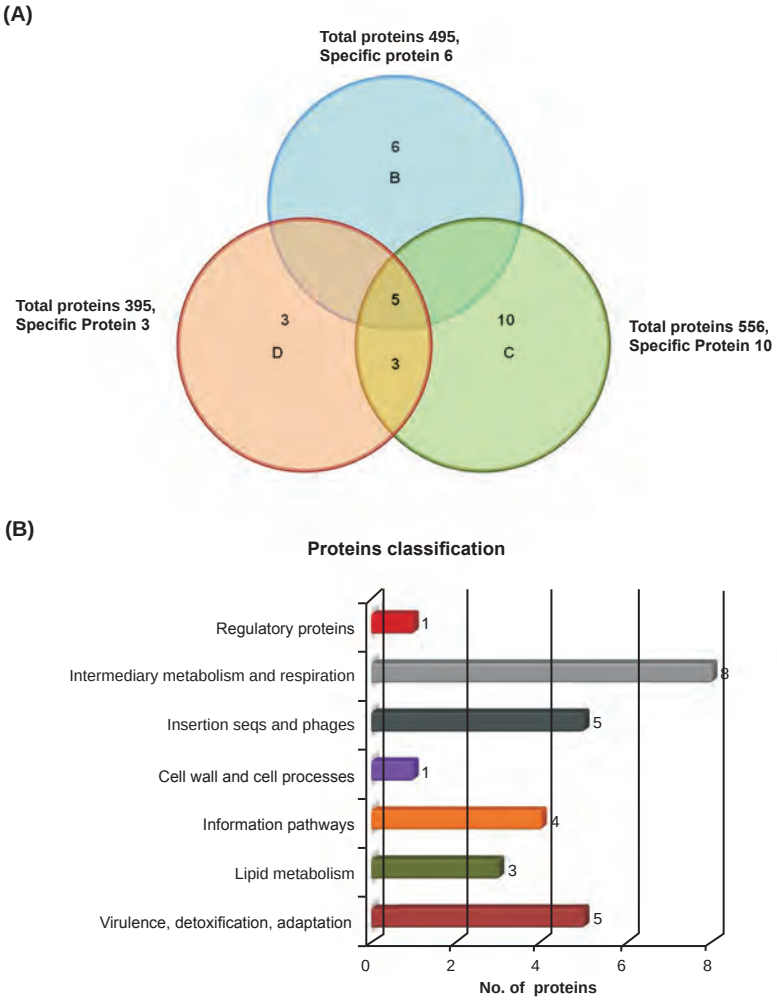

Fig. 2. Upregulated proteins, their distributions, amount of overlap and functional classification in *Mtb* clinical isolates: **(A)** Upregulated proteins distribution and amount of overlap in clinical drug resistant isolates. **(B)** Functional classification of differentially expressed proteins according to the TubercuList Server.

However, in such studies protein (s) that could be modulated *in vivo* during or after acquiring *in vivo* drug resistance are missed out[22]. The ever-increasing evidences suggest that the expression of genes[22] and/or proteins[23] in clinical isolates is markedly different from the laboratory maintained H37Rv strain; suggesting that majority of these observations may not have direct impact in real life scenario.

In the present study, several proteins were identified that were upregulated in drug resistant isolates. Of the 27 upregulated proteins, five were upregulated in all sequential resistant isolates (B, C and D). Approximately half of these overexpressed proteins are reported as essential for *in vitro* growth of *Mtb*[21-24]. Although their functional role in drug resistance is elusive, validation of these proteins as a biomarker of drug resistance will provide a scope for finding an effective candidate drug for MDR-TB. It also needs to be emphasized that identifying a protein as "upregulated" does not necessarily imply that it is a true determinant of drug resistance, because it is possible that some or all of these proteins are associated with adaptability of the *Mtb* to survive longer in the host system.

Among the upregulated proteins, rpoA (Rv3457c) was found 2.9 folds upregulated in the drug resistant isolates. In a similar study compensatory mutations in *rpoA* and *rpoC* (Rv0668) were identified particularly in more than 30 per cent of RIF-resistant strains and the authors proposed that mutation in these alleles could also be associated with MDR[24]. Though we have

Table I. Details of upregulated proteins identified by mass spectrometry in drug resistant isolates

Spot No. †	Proteins identified	Open reading frame (ORF) No.	Accession No.	Mascot score	Nominal mass (Da)	Isoelectric point (pI)	No. of peptides matched	Sequence coverage (%)	Densitometric ratio of protein upregulation between sensitive vs. resistant isolates			Protein classification according to Pasteur Institute of Genomics (TubercuList)
									B	C	D	
1	Aconitase hydratase*	Rv1475c	NP_215991	132	102728	4.95	18	21	1:1.0	1:4.36	1:1.71	7
2	Aconitase hydratase*	Rv1475c	NP_215991	126	102728	4.95	17	21	1:1.22	1:4.70	1:1.58	7
3	Probable fadB protein	Rv0860	NP_215375	171	76170	5.42	24	33	1:1.0	1:1.0	1:2.08	1
4	Chaperonin protein dnaK (HSP70)*	Rv0350	NP_214864	99	66790	4.85	11	22	1:6.88	1:7.88	1:4.76	0
5	Chaperonin protein dnaK (HSP70)*	Rv0350	NP_214864	165	66790	4.85	22	38	1:156.2	1:195.6	1:66.4	0
6	60kDa Chaperonin 2 (cpn60-2, groEL2)*	Rv0440	NP_214954	128	56692	4.85	16	29	1:1.22	1:4.70	1:1.58	0
7	60 kDa Chaperonin 2 (groEL2)*	Rv0440	NP_214954	231	56659	4.85	19	49	1:1.09	1:4.24	1:1.27	0
8	Hypothetical protein	Rv2004c	NP_216520	53	54959	5.89	10	29	1:54.21	1:132.28	1:49.82	5
9	Glutamyl tRNA (Gln) amidotransferase	Rv3011c	NP_217527	98	51787	4.91	9	27	1:2.26	1:1.13	1:0.93	2
10	Elongation factor Tu (EF-Tu)	Rv0685	NP_215199	248	43556	5.28	24	69	1:1.38	1:3.07	1:0.97	2
11	DNA directed RNA polymerase α-chain (rpoA)	Rv3457c	NP_217974	205	37740	4.64	16	43	1:1.7	1:2.98	1:1.0	2
12	Putative phosphoserine aminotransferase	Rv0884c	NP_215399	90	40266	4.77	7	28	1:2.21	1:0.70	1:1.75	7
13	Probable fabG4 protein	Rv0242c	NP_214756	207	46916	6.04	17	52	1:1.01	1:1.0	1:2.08	1
14	Probable moxR protein	Rv1479	YP_177816	103	40738	5.96	12	37	1:1.0	1:2.08	1:1.0	9
15	Probable lipase	Rv2970c	NP_217486	82	43685	6.33	6	30	1:1.19	1:2.28	1:1.0	7
16	Hypothetical protein	Rv1437	NP_217230	88	35519.79	5.14	10	41	1:1.91	1:2.82	1:1.28	7
17	Hypothetical protein	Rv3075c	NP_217591	98	33194	4.73	9	28	1:1.73	1:8.12	1:2.65	5
18	Glyceraldehyde-3-phosphate dehydrogenase	Rv1436	NP_215952	71	36105	5.19	6	17	1:1.0	1:20.61	1:7.0	7
19	Antigen 84 (wag31)*	Rv2145c	NP_216661	120	28260	4.8	8	43	1:12.0	1:7.0	1:3.0	3

Contd...

Spot No. †	Proteins identified	Open reading frame (ORF) No.	Accession No.	Mascot score	Nominal mass (Da)	Isoelectric point (pI)	No. of peptides matched	Sequence coverage (%)	Densitometric ratio of protein upregulation between sensitive vs. resistant isolates			Protein classification according to Pasteur Institute of Genomics (TubercuList)
									B	C	D	
20	Probable quinone oxidoreductase (qor)	Rv1454c	NP_215970	80	34140	5.37	7	42	1:1.15	1:6.42	1:1.89	7
21	Antigen precursor (MPT51)	Rv3803c	YP_178017	64	31069	6.13	5	23	1:2.19	1:1.12	1:1.02	1
22	Ribosome recycling factor (RRF)	Rv2882c	NP_217398	108	20815	5.71	10	67	1:1.0	1:1.0	1:3.45	2
23	Hypothetical protein (GarA)	Rv1827	NP_216343	87	17240	4.29	5	45	1:7.66	1:1.37	1:1.16	5
24	bfrA (Bacterioferritin)	Rv1876	NP_216392	90	18443	4.5	12	82	1:9.27	1:37.5	1:38.58	7
25	Hypothetical protein	Rv2204c	NP_216720	103	12707	4.4	6	51	1:9.11	1:1.0	1:1.04	5
26	Hypothetical protein	Rv0543c	NP_215057	67	14743	5.2	5	38	1:3.03	1:1.01	1:1.53	5
27	10 kDa Chaperonin (cpn 10, groES protein)*	Rv3418c	NP_217935	162	10667	4.62	10	99	1:1.0	1:10.86	1:9.28	0

Cut-off limit ≥2.0 fold for overexpression of proteins. † Spot number of the protein as marked in Fig. 1. *Some proteins are having mobility difference.

Table II. MALDI-TOF/TOF (MS/MS) analysis of overexpressed proteins in MDR-TB clinical isolates

Spot No.†	Peak mass (Da)	Protein identified	Rv No.	Nominal Mass (Da)	pI	Mascot score	Sequence peptide
1	829.4455	Aconitase hydratase	Rv1475c	102728	4.95	39	KSYQIYRL
	1132.6464	Aconitase hydratase		102728	4.95	32	RNGGILQYVLRN
	1170.5177	Aconitase hydratase		102728	4.95	34	RWGQGAFDDFKV
	1299.587	Aconitase hydratase		102728	4.95	37	RIDTPGEADYYRN
2	829.4455	Aconitase hydratase	Rv1475c	102728	4.95	37	KSYQIYRL
	1132.6464	Aconitase hydratase		102728	4.95	39	RNGGILQYVLRN
	1170.5177	Aconitase hydratase		102728	4.95	35	RWGQGAFDDFKV
	1299.587	Aconitase hydratase		102728	4.95	35	RIDTPGEADYYR.N
3	1530.73	Probable fadB protein	Rv0860	76170	5.42	49	KSGSSQPPLQDMIDRM
	1836.9363	Probable fadB protein		76170	5.42	40	KGVDFVIEAVFENQELKH
	2078.1801	Probable fadB protein		76170	5.42	37	KGSQLGLPEVTLGLLPGGGGVTRT
4	1568.0111	Chaperonin protein dnaK (hsp70)	Rv0350	66790	4.85	50	KLLGSFELTGIPPAPRG
	1743.8923	Chaperonin protein dnaK (hsp70)		66790	4.85	37	RATSGDNHLGGDDWDQRV
	2109.2639	Chaperonin protein dnaK (hsp70)		66790	4.85	40	RNGEVLVGQPAKNQAVTNVDRT
5	1062.611	Chaperone protein dnaK	Rv0350	66790	4.85	32	RTTPSIVAFARN
	1226.711	Chaperone protein dnaK		66790	4.85	33	KDAGQIAGLNVLRI
	1645.962	Chaperone protein dnaK		66790	4.85	37	RIVNEPTAAALAYGLDKG
	2613.426	Chaperone protein dnaK		66790	4.85	55	RSETFTTADDNQPSVQIQVYQGERE
6	914.6248	60kDa Chaperonin 2 (cpn60-2, groEL2)	Rv0440	56692	4.85	55	KGRNVVLEKK
	1223.7164	60kDa Chaperonin 2 (cpn60-2, groEL2)		56692	4.85	33	KTIAYDEEARRG
	1266.754	60kDa Chaperonin 2 (cpn60-2, groEL2)		56692	4.85	35	MAKTIAYDEEARR
	1580.0454	60kDa Chaperonin 2 (cpn60-2, groEL2)		56692	4.85	42	REGLRNVAAGANPLGLKR
7	1067.5167	60kDa Chaperonin 2 (cpn60-2, groEL2)	Rv0440	56692	4.85	34	KTIAYDEEARR
	1264.5909	60kDa Chaperonin 2		56692	4.85	36	KEIELEDPYEKI
	1529.7887	60kDa Chaperonin 2		56692	4.85	46	KWGAPTITNDGVSIAKE
	2075.0432	60kDa Chaperonin 2		56692	4.85	66	KTDDVAGDGTTTATVLAQALVRE
8	804.3472	Hypothetical protein	Rv2004c	54959	5.89	35	RERACIRE
	1434.7007	Hypothetical protein		54959	5.89	42	RIEHMVDEFVSGRE
	1796.9013	Hypothetical protein		54959	5.89	45	RIDDAAFLAMDLEFLGRK

Contd...

Spot No.†	Peak mass (Da)	Protein identified	Rv No.	Nominal Mass (Da)	pI	Mascot score	Sequence peptide
9	1094.6241	Glutamyl tRNA (Gln amidotransferase subunit A (gatA)	Rv3011c	51787	4.91	43	RSPYDATLTARL
	1881.9284	gatA		51787	4.91	35	RYGLVACASSLDQGGPCART
	2017.1693	gatA		51787	4.91	36	RQPAALTATVGVKPTYGTVSRY
10	1413.8456	Ef-tu	Rv0685	43556	5.28	32	RQVGVPYILVALNKA
	1693.8359	Ef-tu		43556	5.28	40	RHYAHVDAPGHADYIKN
	2091.0286	Ef-tu		43556	5.28	42	KADAVDDEELLELVEMEVRE
11	1085.5899	DNA directed RNA polymerase α-chain (rpoA)	Rv3457c	37740	4.64	35	KLEVELVVERG
	1485.8157	DNA directed RNA polymerase α-chain (rpoA)		37740	4.64	27	RTLLSSIPGAAVTSIRI
	1611.8113	DNA directed RNA polymerase α-chain (rpoA)		37740	4.64	25	RIDGVLHEFTTVPGVKE
12	1352.6949	Putative phosphoserine aminotransferase		40266	4.77	34	RSLHLTYGEFSAKF
	1900.0249	Putative phosphoserine aminotransferase		40266	4.77	33	MADQLTPHLEIPTAIKPRD
	1928.0345	Putative phosphoserine aminotransferase		40266	4.77	38	RWVPDFLSLPIAVENSLKN
13	1237.69	Probable fabG4 protein	Rv0242c	46916	6.04	44	RQLGVPQPETLRR
	1393.7932	Probable fabG4 protein		46916	6.04	35	RQLGVPQPETLRRY
	1565.862	Probable fabG4 protein		46916	6.04	39	RAGEPPLTGSLLIGGAGRV
14	1425.9539	Probable moxR	Rv1479	40738	5.96	41	KRIIVGQDQLVERM
	1786.142	Probable moxR		40738	5.96	32	RIQFTPDLVPTDIIGTRI
	1983.2737	Probable moxR		40738	5.96	33	RDYVIPQDVIEVIPDVLRH
15	1550.7799	Probable lipase protein	Rv2970c	34146	4.83	34	RVVDLAIDGPAGPIGTRI
	1699.7705	Probable lipase protein		34146	4.83	37	RQHAVGADAIVVSVDYRL
	1740.8621	Probable lipase protein		34146	4.83	22	RIAVAGDSAGGTIAAVIAQRA
16	1315.7195	Phosphoglycerate kinase	Rv1437	42600	4.83	43	RGLLETYHDVLRL
	1420.7409	Phosphoglycerate kinase		42600	4.83	40	KGAFSVVGGGDSAAAVRA
	1683.9461	Phosphoglycerate kinase		42600	4.83	38	RAEGLTGGDILLLENIRF
17	1016.503	Hypothetical protein Rv3075c	Rv3075c	33194	4.73	25	KEFFAEFARD
	1322.619	Hypothetical protein Rv3075c		33194	4.73	60	RWFGDGNADWVRI
	1583.78	Hypothetical protein Rv3075c		33194	4.73	59	RDTGFGEDPATLAYARS

Contd...

Spot No.†	Peak mass (Da)	Protein identified	Rv No.	Nominal Mass (Da)	pI	Mascot score	Sequence peptide
	1648.036	Hypothetical protein Rv3075c		33194	4.73	43	KRLPNVPIVALVETARG
18	1085.7325	Glyceraldehyde-3-phosphate (G-3-P) dehydrogenase	Rv1436	36105	5.19	42	KAIGLVMPQLKG
	1134.7153	G-3-P dehydrogenase		36105	5.19	37	KVLDDEFGIVKG
	1384.8723	G-3-P dehydrogenase		36105	5.19	34	RAAALNIVPTSTGAAKA
	1085.7325	Glyceraldehyde-3-phosphate (G-3-P) dehydrogenase		36105	5.19	42	KAIGLVMPQLKG
19	1088.5844	ag84/wag31	Rv2145c	28260	4.8	22	RLIEENSDLRQ
	1171.6518	ag84/wag31		28260	4.8	31	RANAEQILGEARH
	1413.7358	ag84/wag31		28260	4.8	36	KHSEIMGTINQQRA
20	1011.5087	qor	Rv1454c	34140	5.37	43	RTGEEFSWRA
	1615.8965	qor		34140	5.37	40	KAEAIGVNFIDTYFRS
	2195.0742	qor		34140	5.37	34	KDAGADVVLDYPEDAWQFAGRV
21	1037.4687	Antigen precursor (MPT51)	Rv3803c	31069	6.13	34	RMFYNQYRS
	2044.93	Antigen precursor (MPT51)		31069	6.13	42	KWHDPWVHASLLAQNNTRV
	2132.9473	Antigen precursor (MPT51)		31069	6.13	45	KQWDTFLSAELPDWLAANRG
22	1655.8488	Ribosome recycling factor (RRF)	Rv2882c	20815	5.71	29	RNSDLGVNPTNDGALIRV
	1674.88	Ribosome recycling factor (RRF)		20815	5.71	37	KTTHQYVTQIDELVKH
	2146.0955	Ribosome recycling factor (RRF)		20815	5.71	38	KDLDKTTHQYVTQIDELVKH
23	1291.7384	Hypothetical protein (GarA)	Rv1827	17240	4.29	21	RFLLDQAITSAGRH
	1715.8809	Hypothetical protein (GarA)		17240	4.29	43	RHPDSDIFLDDVTVSRR
	1840.9915	Hypothetical protein (GarA)		17240	4.29	28	REPVDSAVLANGDEVQIGKF
24	1046.5234	Bfr (Bacterioferritin)	Rv1876	18443	4.5	40	MQGDPDVLRL
	1414.7907	Bfr (Bacterioferritin)		18443	4.5	37	RILLLDGLPNYQRI
	1935.8135	Bfr (Bacterioferritin)		18443	4.5	29	RAESFDEMRHAEEITDRI
25	1089.5305	Hypothetical protein Rv2204c	Rv2204c	12707	4.4	22	RYNLFFDDRT
	1281.7399	Hypothetical protein Rv2204c		12707	4.4	28	KTHGVILTEAAAAKA
	1198.665	Hypothetical protein Rv2204c		12707	4.4	35	RIAVQPGGCAGLRY
26	965.4423	Hypothetical protein Rv0543c	Rv0543c	14743	5.2	26	RDDAPYWAKY

Contd...

Spot No.†	Peak mass (Da)	Protein identified	Rv No.	Nominal Mass (Da)	pI	Mascot score	Sequence peptide
	1218.6172	Hypothetical protein Rv0543c		14743	5.2	25	MSVELTQEVSARL
	1595.7916	Hypothetical protein Rv0543c		14743	5.2	19	RLTSDLYGWLTTVARS
27	1034.4976	10kDa Chaperonin groES	Rv3418c	10798	4.62	40	RWDEDGEKRI
	1523.8874	10kDa Chaperonin groES		10798	4.62	21	KEKPQEGTVVAVGPGRW
	1776.0298	10kDa Chaperonin groES		10798	4.62	29	KRIPLDVAEGDTVIYSKY

†Spot number of the proteins marked as in Fig. 1.

not screened our resistant isolate for these mutations, this study supports our proteomic approach and identification of an upregulated rpoA protein in drug resistant isolates. Dussurget *et al*[25] reported that in *M. smegmatis*, IdeR negatively controls iron-uptake and expression of BfrA and BfrB. In our study, BfrA was upregulated in drug resistant isolates, suggesting a role of this protein in inducing resistance to INH. However, such conclusions could not be validated by deletion mutant of *bfrA* and *bfrB* in *Mtb*[22].

We found that the identified proteins spots (Rv1475c, dnaK, groEL2, groES and wag31) had different electrophoretic mobilities in resistant and sensitive isolates. Similar observations have been reported by Mattow *et al*[21] when the protein profile of intra-phagosomal *Mtb* H37Rv was analyzed. It has also been suggested that this shift is determined by the protein modifications rendered during sample preparation and growth conditions adopted by various laboratories[8]. However, this may be the unlikely factor in our study, as we strictly followed the same protocol throughout the study. Further, our analysis was stringent and we considered a particular protein to be "upregulated" only if the observation was consistent in three independent experiments. Even then, we were cautious not to conclude whether the changes in protein mobility necessarily reflect protein modification. Further studies are warranted to establish the role of post-translation modifications in these proteins during drug resistance.

Antigen 84 (Wag31 / Rv2145c) has been demonstrated to be involved in the regulation of cell morphology and pknG mediates the survival of *Mtb* inside the host macrophages, and genome-wide transcriptional analysis reveals that the *pknG* is upregulated during the exposure of INH[26]. Further, *pknG* is associated with the intrinsic resistance of *Mtb* to various anti-TB drugs[26]. Wag31 protein has been found to be overexpressed and involved in regulation of cell morphology. It also plays important role in survival of mycobacteria under oxidative stress[27] and provides optimal substrate for pknA and pknB. In addition, mRNA expression of wag31 was increased by 15.7 folds during the INH and RIF exposure. Earlier also, wag31 has been reported to be over expressed in the MDR isolates[8], supporting our hypothesis that wag31 plays an important role in drug resistance.

Of the five hypothetical proteins (Rv2004c, Rv1437, Rv3075c, Rv2204c and Rv1827) overexpressed in MDR isolates, three (Rv2004c, Rv2204c and Rv1437) could not be assigned to any function in survival or pathogenesis of the bacteria, though Rv3075c has been reported to be overexpressed in streptomycin resistant isolates[11]. In our study, Rv1827 (GarA) was found 7.6 folds upregulated in drug resistant isolates as compared to that in the susceptible isolate A. It has been identified as an optimal substrate for PknB and PknG. The protein is also reported to act as a phosphorylation-dependent molecular switch in mycobacterial signalling process mediated by protein kinases[28-30]. Further, GarA has been found to be predominantly expressed under the exponential growth phase and has been suggested as a regulatory model for glycogen degradation and glutamate metabolism[28-30]. To infer whether the protein overexpression of Rv1827 in drug resistant isolates facilitated increased glycogen accumulation, we quantified the glycogen content of drug resistant and sensitive isolates. The findings were consistent with this hypothesis and the glycogen content was relatively

Table III. Isolates specific proteins overexpression in MDR-TB isolates

S. No.	Spot No.	Proteins identified	Open reading frame (ORF) No.	Accession No.	Mascot score	Nominal mass (Da)	pI	No. of peptides matched	Sequence coverage (%)	Densitometric ratio of protein expression between sensitive vs. resistant isolates			Protein classification according to Pasteur Institute of Genomics (TubercuList)
										B	C	D	
Proteins overexpressed in MDR isolate (Isolate B)													
1	9	Glutamyl tRNA (Gln amidotransferase	Rv3011c	NP_217527	98	51787	4.91	9	27	1:2.26	1:1.13	1:0.93	2
2	12	Putative phosphoserine aminotransferase	Rv0884c	NP_215399	90	40266	4.77	7	28	1:2.21	1:0.70	1:1.75	7
3	21	Antigen precursor (MPT51)	Rv3803c	YP_178017	64	31069	6.13	5	23	1:2.19	1:1.12	1:1.02	1
4	23	Hypothetical protein (GarA)	Rv1827	NP_216343	87	17240	4.29	5	45	1:7.66	1:1.37	1:1.16	5
5	25	Hypothetical protein	Rv2204c	NP_216720	103	12707	4.4	6	51	1:9.11	1:1.0	1:1.04	5
6	26	Hypothetical protein	Rv0543c	NP_215057	67	14743	5.2	5	38	1:3.03	1:1.01	1:1.53	5
Proteins overexpressed in MDR isolate (Isolate C)													
1	1	Aconitase hydratase	Rv1475c	NP_215991	132	102728	4.95	18	21	1:1.0	1:4.36	1:1.71	7
2	2	Aconitase hydratase	Rv1475c	NP_215991	126	102728	4.95	17	21	1:1.22	1:4.70	1:1.58	7
3	6	60kDa Chaperonin 2 (cpn60-2, groEL2)	Rv0440	NP_214954	128	56692	4.85	16	29	1:1.22	1:4.70	1:1.58	0
4	7	60 kDa Chaperonin 2 (groEL2)	Rv0440	NP_214954	231	56659	4.85	19	49	1:1.09	1:4.24	1:1.27	0
5	10	Elongation factor Tu (EF-Tu)	Rv0685	NP_215199	248	43556	5.28	24	69	1:1.38	1:3.07	1:0.97	2
6	11	DNA directed RNA polymerase α-chain (rpoA)	Rv3457c	NP_217974	205	37740	4.64	16	43	1:1.7	1:2.98	1:1.0	2

Contd...

S. No.	Spot No.	Proteins identified	Open reading frame (ORF) No.	Accession No.	Mascot score	Nominal mass (Da)	pI	No. of peptides matched	Sequence coverage (%)	Densitometric ratio of protein expression between sensitive vs. resistant isolates			Protein classification according to Pasteur Institute of Genomics (TubercuList)
										B	C	D	
7	14	Probable moxR protein	Rv1479	YP_177816	103	40738	5.96	12	37	1:1.0	1:2.08	1:1.0	9
8	15	Probable lipase	Rv2970c	NP_217486	82	43685	6.33	6	30	1:1.19	1:2.28	1:1.0	7
9	16	Hypothetical protein	Rv1437	NP_217230	88	35519.79	5.14	10	41	1:1.91	1:2.82	1:1.28	7
10	20	Probable quinone oxidoreductase (qor)	Rv1454c	NP_215970	80	34140	5.37	7	42	1:1.15	1:6.42	1:1.89	7
Proteins overexpressed in MDR isolate (Isolate D)													
1	3	Probable fadB protein	Rv0860	NP_215375	171	76170	5.42	24	33	1:1.0	1:1.0	1:2.08	1
2	13	Probable fabG4 protein	Rv0242c	NP_214756	207	46916	6.04	17	52	1:1.01	1:1.0	1:2.08	1
3	22	Ribosome recycling factor (RRF)	Rv2882c	NP_217398	108	20815	5.71	10	67	1:1.0	1:1.0	1:3.45	2
Proteins overexpressed only in MDR isolates (Isolates C and D)													
1	17	Hypothetical protein	Rv3075c	NP_217591	98	33194	4.73	9	28	1:1.73	1:8.12	1:2.65	5
2	18	Glyceraldehyde-3-phosphate	Rv1436	NP_215952	71	36105	5.19	6	17	1:1.0	1:20.61	1:7.0	7
3	27	10 kDa chaperonin (cpn 10, groES)	Rv3418c	NP_217935	162	10667	4.62	10	99	1:1.0	1:10.86	1:9.28	0
Proteins overexpressed only in all MDR isolates (Isolates B, C and D)													
1	4	Chaperonin protein dnaK (HSP70)	Rv0350	NP_214864	99	66790	4.85	11	22	1:6.88	1:7.88	1:4.76	0
2	5	Chaperonin protein dnaK (HSP70)	Rv0350	NP_214864	165	66790	4.85	22	38	1:156.2	1:195.6	1:66.4	0
3	8	Hypothetical protein	Rv2004c	NP_216520	53	54959	5.89	10	29	1:54.21	1:132.28	1:49.82	5
4	19	Antigen 84 (wag31)	Rv2145c	NP_216661	120	28260	4.8	8	43	1:12.0	1:7.0	1:3.0	3
5	24	bfrA (Bacterioferritin)	Rv1876	NP_216392	90	18443	4.5	12	82	1:9.27	1:37.5	1:38.58	7

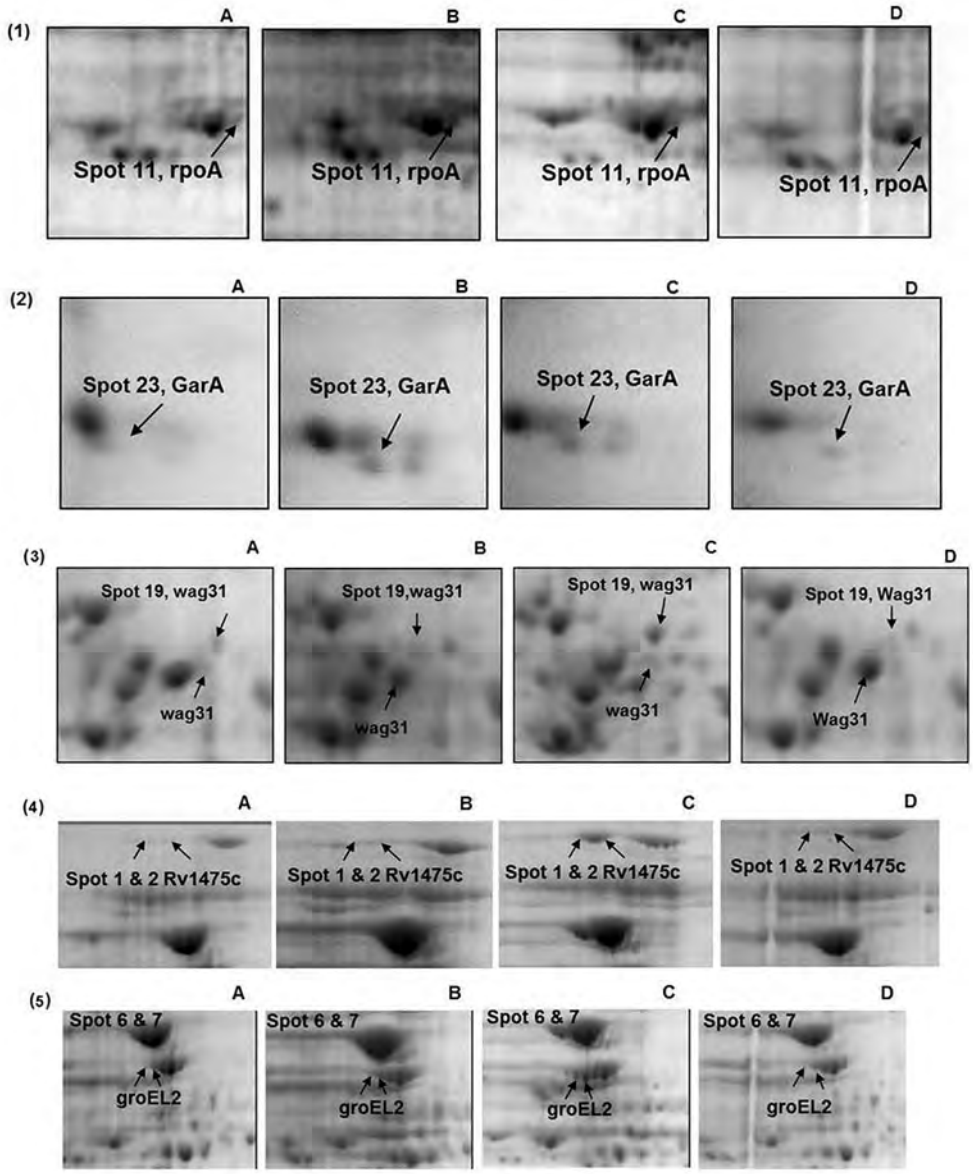

Fig. 3. Magnified region of overexpressed proteins present in MDR isolates. Zoomed in regions of 2D-gel images showing the overexpressed proteins in drug resistant isolates: (1) rpoA (2) GarA (3) wag31 (4) Rv1475c (5) groEL2.

more in resistant isolates (B, C and D) than the sensitive isolate A. However, the difference in accumulation was significant up to 7th day, but not after 15th day of growth, suggesting that the expression of Rv1827 might be an important marker of dormancy. While the precise role of glycogen storage in mycobacteria is not known, glycogen stores may serve as a reservoir of carbon and energy that can be mobilized by mycobacteria for survival during periods of carbon starvation. Such observations have been made in other bacteria such as *Vibrio cholerae* during transition stage between host and aquatic environments[31].

In our study, the identified proteins cannot be categorically classified as drug resistance-specific alterations, as this can also happen due to host immune response or stresses encountered by the individual strain *in vivo*. Most of these proteins are essential for the survival of mycobacteria in phagosome or as virulence factor (unpublished observation), but their role in drug resistance cannot be ruled out completely.

The 2D gel electrophoresis followed by MS-based proteomic analysis on sequential isolates showed approximately 500 proteins per gel. This resolution is

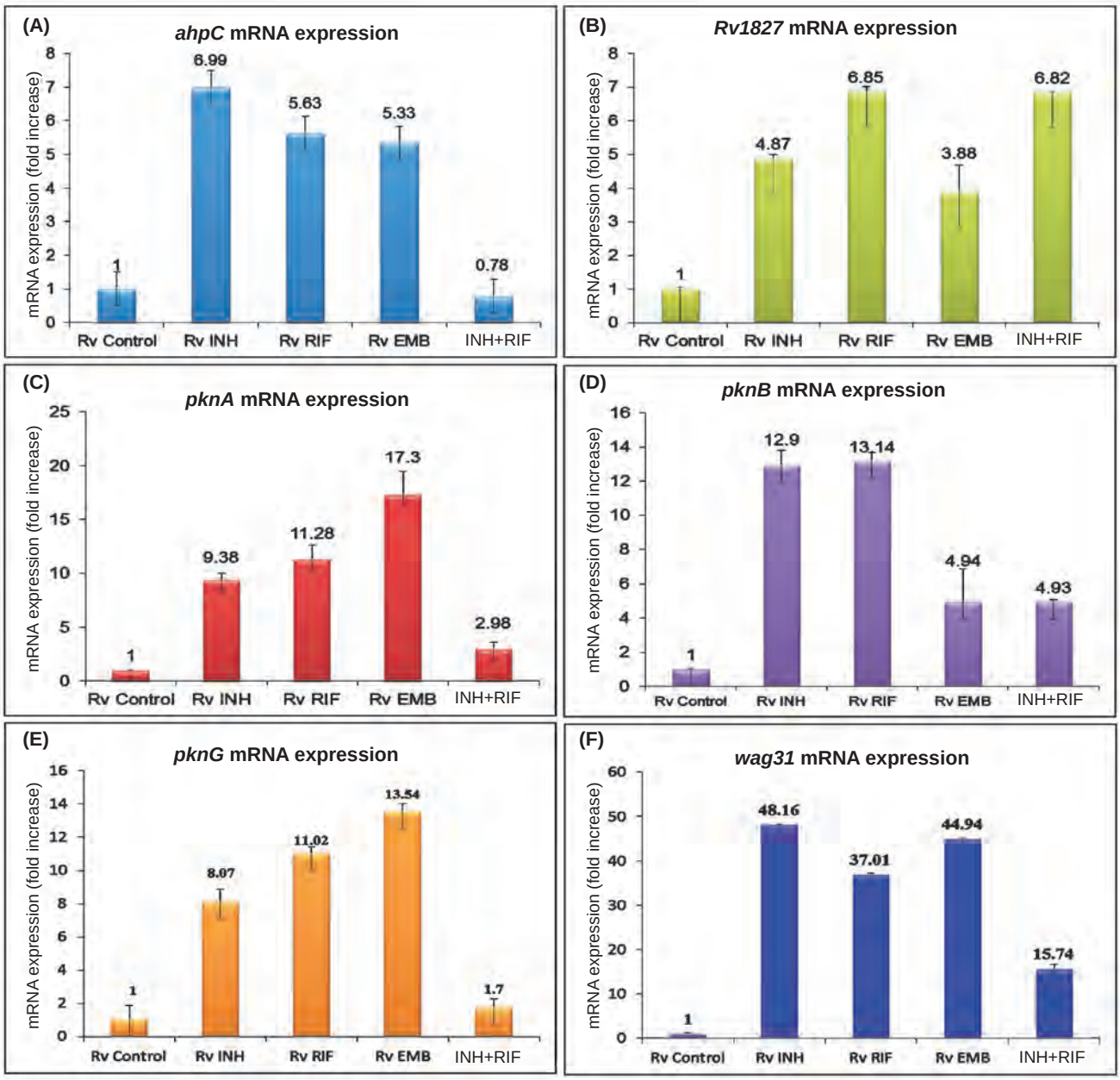

Fig. 4. Detection of mRNA expression in drug treated *Mtb* isolates. Isolates treated with isoniazid (INH, 0.1µg/ml), rifampicin (RIF, 1.0µg/ml), ethambutol (EMB, 5µg/ml) and INH (0.1µg/ml) +RIF (1.0µg/ml). The expression is represented in fold increase. The *16sRNA* was used as internal control for mRNA expression analysis. mRNA expression of **(A)** *ahpC*, **(B)** *Rv1827*, **(C)** *pknA*, **(D)** *pknB*, **(E)** *pknG* and **(F)** *wag31*. Calculation and statistical analysis carried out by using the $2^{-\Delta\Delta C}{}_{T}$ method and results presented in fold increase.

much less than expected, as there are 4000 predicted genes in *Mtb*. The poor sensitivity observed in the present study could be attributed to various reasons such as the extraction protocol; low resolution power of the coomassie brilliant blue stain used or due to the IEF-strips (*p*H 4-7). These strips resolve only the proteins having isoelectric point in the range of 4-7.

In conclusion, our study highlights the intricacies associated with sequential clinical isolates of *Mtb*, a rare opportunity for any laboratory, which is a natural phenomenon and cannot be generated artificially in the laboratory. The sequential isolation of four isolates from the same patient during the treatment period showed a phenomenon where a sensitive isolate turned

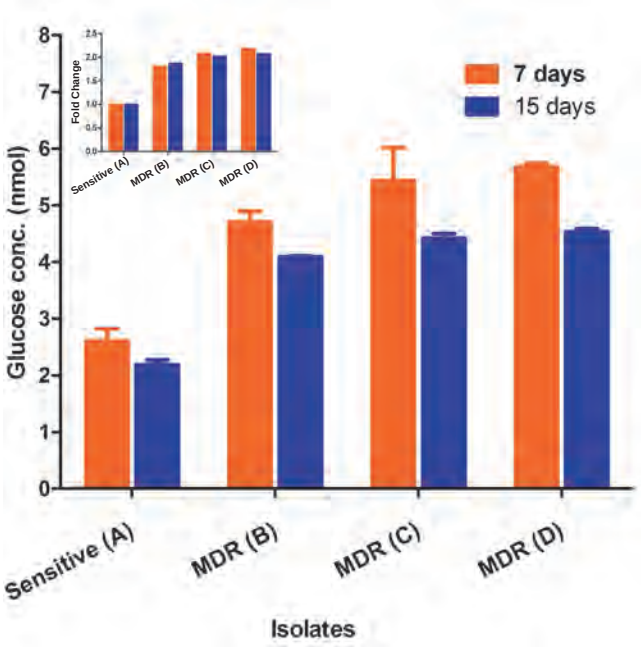

Fig. 5. Glycogen content of sensitive and MDR isolates of *Mtb*. Glycogen was extracted after 7 and 15 days and glycogen storage was determined by the estimation of glucose. The bars indicate glucose concentration of *Mtb* cells after 7 and 15 days, respectively. The inset shows the glycogen content in terms of fold increase after normalization in comparison to sensitive *vs.* MDR TB clinical isolates. Values are mean ± SD (n=3).

to a multidrug resistant isolate. It is possible that some of the upregulated proteins identified from MDR clinical isolates of *Mtb* in the present study may prove as potential biomarkers of drug resistance in future.

Acknowledgment

This work was supported by grant from Department of Biotechnology, Government of India to the last author (SS) (BT/ PR7880/ MED/ 14/1161/2006). The first author (AS) received financial assistance in the form of senior research fellowship from the Indian Council of Medical Research, Government of India. The authors thank Shri Sandeep Kadian for his help in spoligotyping data analysis and Dr V.M. Katoch, Director-General of Indian Council of Medical Research, New Delhi, for allowing us to use his laboratory facilities for MALDI-TOF/TOF MS at JALMA, Agra.

Patent Applied: Sarman Singh, Gopinath K., Amit Singh and Niti Singh. Novel protein markers of drug resistance in *Mycobacterium tuberculosis* (1752/DEL/2008).

References

1. World Health Organization. *Global Tuberculosis Report 2014*. Available from: *http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf?ua=1*, accessed on November 3, 2014.
2. Rattan A, Kalia A, Ahmad N. Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives. *Emerg Infect Dis* 1998; *4* : 195-209.
3. Ramaswamy SV, Reich R, Dou SJ, Jasperse L, Pan X, Wanger A, *et al*. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 2003; *47* : 1241-50.
4. Gryadunov D, Mikhailovich V, Lapa S, Roudinskii N, Donnikov M, Panikov S, *et al*. Evaluation of hybridization on oligonucleotide microarrays for analysis of drug-resistant *Mycobacterium tuberculosis*. *Clin Microbiol Infect* 2005; *11* : 531-9.
5. Louw GE, Warren RM, Gey van Pittius NC, McEvoy CR,Van Helden PD, Victor TC. A balancing act: efflux/influx in mycobacterial drug resistance. *Antimicrob Agents Chemother* 2009; *53* : 3181-9.
6. Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, *et al*. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc Natl Acad Sci USA* 1999; *96* : 12833-8.
7. Wang R, Marcotte EM. The proteomic response of *Mycobacterium smegmatis* to anti-tuberculosis drugs suggests targeted pathways. *J Proteome Res* 2008; *7* : 855-65.
8. Jiang X, Zhang W, Gao F, Huang Y, Lu C, Wang H. Comparison of the proteome of isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis*. *Microb Drug Resist* 2006; *12* : 231-8.
9. Sharma P, Kumar B, Singhal N, Katoch VM, Venkatesan K, Chauhan DS, *et al*. Streptomycin induced protein expression analysis in *Mycobacterium tuberculosis* by two-dimensional gel electrophoresis & mass spectrometry. *Indian J Med Res* 2010; *132* : 400-8.
10. Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, *et al*. Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. *Mol Microbiol* 1999; *33* : 1103-17.
11. Sharma P, Kumar B, Gupta Y, Singhal N, Katoch VM, Venkatesan K, *et al*. Proteomic analysis of streptomycin resistant and sensitive clinical isolates of *Mycobacterium tuberculosis*. *Proteome Sci* 2010; *8* : 59.
12. Gopinath K, Singh S. Multiplex PCR assay for simultaneous detection and differentiation of *Mycobacterium tuberculosis*, *Mycobacterium avium* complexes and other Mycobacterial species directly from clinical specimens. *J Appl Microbiol* 2009; *107* : 425-35.
13. Singh A, Gopinath K, Singh N, Singh S. Deciphering the sequential events during *in vivo* acquisition of drug resistance in *Mycobacterium tuberculosis*. *Int J Mycobacteriol* 2014; *3* : 36-40.
14. Patricia TK, Cubica GP. *Public health mycobacteriology*: *a guide for the level III laboratory*. Atlanta, Ga: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control; 1995.
15. Demay C, Liens B, Burguière T, Hill V, Couvin D, Millet J, *et al*. SITVITWEB - a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect Genet Evol* 2012; *12* : 755-66.

16. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. *Anal Biochem* 1976; *72* : 248-54.

17. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 2006; *1* : 2856-60.

18. Schulze U, Larsen ME, Villadsen J. Determination of intracellular trehalose and glycogen in *Saccharomyces cerevisiae*. *Anal Biochem* 1995; *228* : 143-9.

19. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}$T method. *Methods* 2001; *25* : 402-8.

20. Reddy PV, Puri RV, Khera A, Tyagi AK. Iron storage proteins are essential for the survival and pathogenesis of *Mycobacterium tuberculosis* in THP-1 macrophages and the guinea pig model of infection. *J Bacteriol* 2012; *194* : 567-75.

21. Mattow J, Siejak F, Hagens K, Becher D, Albrecht D, Krah A, *et al*. Proteins unique to intraphagosomally grown *Mycobacterium tuberculosis*. *Proteomics* 2006; *6* : 2485-94.

22. Homolka S, Niemann S, Russell DG, Rohde KH. Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog* 2010; *6* : e1000988.

23. Pheiffer C, Betts JC, Flynn HR, Lukey PT, van Helden P. Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv. *Microbiology* 2005; *151* : 1139-50.

24. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, *et al*. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 2011; *44* : 106-10.

25. Dussurget O, Rodriguez M, Smith I. Protective role of the *Mycobacterium smegmatis* IdeR against reactive oxygen species and isoniazid toxicity. *Tuber Lung Dis* 1998; *79* : 99-106.

26. Wolff KA, Nguyen HT, Cartabuke RH, Singh A, Ogwang S, Nguyen L, *et al*. Protein kinase G is required for intrinsic antibiotic resistance in mycobacteria. *Antimicrob Agents Chemother* 2009; *53* : 3515-9.

27. Mukherjee P, Sureka K, Datta P, Hossain T, Barik S, Das KP, *et al*. Novel role of Wag31 in protection of mycobacteria under oxidative stress. *Mol Microbiol* 2009; *73* : 103-19.

28. O'Hare HM, Durán R, Cerveñansky C, Bellinzoni M, Wehenkel AM, Pritsch O, *et al*. Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* 2008; *70* : 1408-23.

29. Kang CM, Abbott DW, Park ST, Dascher CC, Cantley LC, Husson RN. The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev* 2005; *19* : 1692-704.

30. England P, Wehenkel A, Martins S, Hoos S, André-Leroux G, Villarino A, *et al*. The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. *FEBS Lett* 2009; *583* : 301-7.

31. Bourassa L, Camilli A. Glycogen contributes to the environmental persistence and transmission of *Vibrio cholerae*. *Mol Microbiol* 2009; *72* : 124-38.

Reprint requests: Dr Sarman Singh, Division of Clinical Microbiology & Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110 029, India
e-mail: sarman_singh@yahoo.com, sarman.singh@gmail.com