BMC Infectious Diseases

Research article

Genotyping and drug resistance patterns of *M. tuberculosis* strains in Pakistan

Mahnaz Tanveer[1], Zahra Hasan[1], Amna R Siddiqui[2], Asho Ali[1], Akbar Kanji[1], Solomon Ghebremicheal[3] and Rumina Hasan*[1]

Address: [1]Department of Pathology and Microbiology, The Aga Khan University, Stadium Road Karachi, Pakistan, [2]Department of Community Health Sciences, The Aga Khan University, Stadium Road Karachi, Pakistan and [3]Department of Bacteriology, Swedish Institute for Infectious Diseases Control, Stockholm, Sweden

Email: Mahnaz Tanveer - mahnaz.tanveer@aku.edu; Zahra Hasan - maqboola.dojki@aku.edu; Amna R Siddiqui - rehana.siddiqui@aku.edu; Asho Ali - asho.ali@aku.edu; Akbar Kanji - akbar.kanji@aku.edu; Solomon Ghebremicheal - solomon.ghebremicheal@smi.ki.se; Rumina Hasan* - rumina.hasan@gmail.com

* Corresponding author

Abstract

Background: The incidence of tuberculosis in Pakistan is 181/100,000 population. However, information about transmission and geographical prevalence of *Mycobacterium tuberculosis* strains and their evolutionary genetics as well as drug resistance remains limited. Our objective was to determine the clonal composition, evolutionary genetics and drug resistance of *M. tuberculosis* isolates from different regions of the country.

Methods: *M. tuberculosis* strains isolated (2003–2005) from specimens submitted to the laboratory through collection units nationwide were included. Drug susceptibility was performed and strains were spoligotyped.

Results: Of 926 *M. tuberculosis* strains studied, 721(78%) were grouped into 59 "shared types", while 205 (22%) were identified as "Orphan" spoligotypes. Amongst the predominant genotypes 61% were Central Asian strains (CAS ; including CAS1, CAS sub-families and Orphan Pak clusters), 4% East African-Indian (EAI), 3% Beijing, 2% poorly defined TB strains (T), 2% Haarlem and LAM (0.2). Also TbD1 analysis (*M. tuberculosis* specific deletion 1) confirmed that CAS1 was of "modern" origin while EAI isolates belonged to "ancestral" strain types.

Prevalence of CAS1 clade was significantly higher in Punjab (P < 0.01, Pearsons Chi-square test) as compared with Sindh, North West Frontier Province and Balochistan provinces. Forty six percent of isolates were sensitive to five first line antibiotics tested, 45% were Rifampicin resistant, 50% isoniazid resistant. MDR was significantly associated with Beijing strains (P = 0.01, Pearsons Chi-square test) and EAI (P = 0.001, Pearsons Chi-square test), but not with CAS family.

Conclusion: Our results show variation of prevalent *M. tuberculosis* strain with greater association of CAS1 with the Punjab province. The fact that the prevalent CAS genotype was not associated with drug resistance is encouraging. It further suggests a more effective treatment and control programme should be successful in reducing the tuberculosis burden in Pakistan.

Background

Tuberculosis (TB) remains a major cause of morbidity and mortality worldwide, causing more than 2 million deaths a year [1,2]. Pakistan with a population of 140 million and a growth rate of 3.5% [3] has the seventh highest tuberculosis rate despite widespread BCG vaccination. TB prevalence in Pakistan can be attributed to poor health care systems and limited diagnostic and treatment modalities for TB[4]. The TB problem is also compounded by multi-drug resistance (MDR, resistance to at least rifampicin and isoniazid), WHO report suggests that globally 3% of *M. tuberculosis* isolates are MDR-TB[5]. Resistance to TB drugs is recognized in Pakistan[6,7]. While community based information is lacking, laboratory data suggests an increasing frequency of MDR from 14% in 1999 to 28% in 2004[6] and 47% in 2006[8].

International databases such as the SpolDB4.0 have revealed the clonal structure of *M. tuberculosis* isolates in different geographical settings. SpolDB4.0 data base further defines super families specific to certain locations[9]. Genotypic information has further expanded our understanding of strain prevalence and transmission [10-13]. A few predominant genotypes circulating throughout the world e.g. Beijing, Haarlem, and African clusters have been associated with a number of major outbreaks [14-16]. These major strain groups have been described as being predominant pathotypes in the world [17]. The abundance of polymorphism indicates that transposition and homologous recombination are the major events contributing to the diversity of *M. tuberculosis* strains [18]. In addition, polymorphism seen with different molecular markers also describes mutual association. This supports the hypothesis that *M. tuberculosis* has a strong clonal population structure [18]. In support of phylogeographical population structure of *M. tuberculosis*, differences in strain genetics may be responsible for the variation in BCG efficacy [19-22].

Predominant *M. tuberculosis* clades from the Indian subcontinent include Central Asian strain (CAS) [9,23] and Beijing strains [11,24-27]. Central Asian strain 1 (CAS1) are defined by absence of spacers 4–7 and 23–34 [28]. While, Beijing strains were characterized with the absence of 1–34 spacers in direct repeat region (DR). Beijing strains are reported to constitute about 50% of strains in far East-Asia and 13% of isolates globally[29]. East African-Indian strains, the T clade and Haarlem strains have also reported from India, Afghanistan and Iran[30,31]. In Pakistan predominance of CAS1 (39%) with a 6% prevalence of Beijing isolates has previously been reported [32].

Globally, MDR-TB outbreaks have been associated with Beijing and Haarlem families [33,34]. In order to understand the population structure of *M. tuberculosis* in Paki-

stan, strains from the four provinces, Punjab, Sindh, Balochistan and NWFP were spoligotyped. Genotypic information was correlated with drug resistance to determine association between strain types and MDR. Predominant clades obtained were further analyzed to distinguish between "ancestral" versus "modern" lineages of tubercle bacilli based on the presence or absence of the TbD1 region.

Methods

Mycobacterial strain collection

This study was conducted on *M. tuberculosis* strains isolated at the Aga Khan University Hospital (AKUH) in Karachi during the 3 year period 2003–2005. Specimens were from collection points situated in all four provinces of Pakistan. AKUH is a tertiary care hospital in Karachi and its clinical microbiology laboratory receives specimens through more than 100 collection points situated in all four provinces of Pakistan. All samples were delivered within a period of 24 hrs of collection and were processed for culture and sensitivity testing. Strains were stored at -70°c in 15% glycerol phosphate broth. During the study period 6,067 samples were processed from which 2208 strains were isolated. Based on the stratified random sampling method[32], a total of 926 strains were included in this study (It was ensured that only one sample per individual is included in our analysis). The largest 234 (25%) were from 14 different locations in Karachi. A further 691 strains were from the 4 provinces of Pakistan (excluding Karachi); 256 from Punjab, 224 from Sindh, 207 from the North West Frontier Province (NWFP) and 5 from Balochistan. We were not able to classify patients on basis of prior therapy since treatment history was not available. The *M. tuberculosis* isolates studied included both pulmonary (n = 850) and extra-pulmonary (n = 76) samples.

Microbiological methods

Mycobacterial cultures and antibiotic susceptibility testing

Mycobacterial cultures were performed on liquid as well as solid media. Respiratory samples were decontaminated using N-acetyl-L-cysteine (NALC) sodium hydroxide prior to culture. Samples from sterile sites were processed without decontamination[35]. All specimens were concentrated by centrifugation (3000 × g) for 30 minutes and sediments cultured at 37°C using BACTEC 460 (Becton Dickinson Diagnostic Instruments Systems) and Lowenstein Jensen (LJ) medium. The growth index of inoculated BACTEC vials was checked for four weeks, LJ slants were incubated for up to 8 weeks. *M. tuberculosis* was identified by BACTEC NAP TB differentiation test (Becton Dickinson, USA).

Susceptibility testing was performed using standard agar proportion method on enriched Middle brook 7H10 medium (BBL) at the following final drug concentrations;

rifampicin 1 ug/ml and 5 ug/ml, isoniazid 0.2 ug/ml and 1 ug/ml, streptomycin 2 ug/ml and 10 ug/ml and ethambutol 5 ug/ml and 10 ug/ml, ethionamide 5 ug/ml, capreomycin 10 ug/ml, cycloserine 30 ug/ml and ciprofloxacin 2 ug/ml. [36-38]. Pyrazinamide sensitivity was carried out using BACTEC 7H12 medium pH6.0 at 100 ug/ml (BACTEC™ PZA test medium, Becton Dickinson USA) in accordance with manufacturers instructions. To ensure selection of high level resistance strains for purposes of this study however, only resistance to the higher concentrations were used for analysis. Multidrug resistance (MDR) was defined as resistance to at least isoniazid and rifampicin.

Molecular methods

Mycobacteria were cultured on 7H10 Middle brook agar. DNA extraction was carried out from mycobacterial colonies using the CTAB method[39]. Spoligotyping was carried out using a commercially available kit from Isogen Bioscience BV, Maarssen, The Netherlands according to the manufacturer's instructions. Spoligotyping based on the 43 spacers of the DR region of *M. tuberculosis* complex was carried out using primers DRa 5'GGTTTTGGGT CTGACGAC3' and DRb 5'CCGAGAGGGGACGGAAAC 3' as originally described by Kamerbeek et al[24].

TbD1 analysis

TbD1; "*M. tuberculosis* specific deletion 1" consists of two genes encoding membrane protein (mmpS6 and mmpL6). TbD1 was originally identified as a 2153-bp fragment[40]. TbD1 deletion analysis was done by PCR as described by Brosch et al[41]. Two isolates from each cluster of CAS1 and its sub-families were tested for both primers complementary to flanking sequences and internal sequences in order to confirm deleted region (TbD1) in our Mycobacterium isolates. Sequences inside or flanking regions were obtained from the website http://www.pnas.org/cgi/data/052548299/DC1/1/[11].

Data analysis

Spoligotyping results were entered in the Bionumerics Software, Applied Maths Program, BioSystematica, UK. Dendrograms were generated using the unweighted pair-group method with arithmetic averages (UPGMA) calculation. A cluster (shared types) was defined as two or more isolates from different patients with identical spoligotype patterns, whereas, non-clustered strains had 'orphan' spoligotype patterns. The spoligotypes were compared with the most prevalent *M. tuberculosis* subfamilies as identified by the World Spoligotyping Database SpolDB4.0 of Pasteur Institute of Guadeloupe http://www.pasteur-guadeloupe.fr/tb/bd_myco.html[42]. The SpolDB4.0 information system is an automated Access-based labeling and matching system for spoligotyping. SpolDB4.0 which included ~40,000 isolates split into 1,939 shared types

and ~3,530 orphan profiles. We compared each of our clusters with the shared types (STs) present in SpolDB4.0 [29].

Pearson's Chi-squared test was used to determine statistical association between strain types and specific parameters, Statistical Package for Social Science Software (SPSS, USA) was used for analysis. P values < 0.05 were considered significant. Considering that the isolates were independent the association between cluster types and drug resistance was analysed by the use of multilogistic regression model. The magnitude of the association was estimated by using the odds ratio (OR) and 95% confidence interval (CIs)

Results

Study population

A total of nine hundred and twenty six *M. tuberculosis* clinical isolates were studied. These were from pulmonary (n = 850) and extra-pulmonary (n = 76) sources. Of the pulmonary isolates studied, 411 (48%) were from the province of Sindh of which 195 (23%) were from Karachi; 232 (27%) were from the province of Punjab; 202(22%) from the North West Frontier Province (NWFP) and 5 (0.4%) from Balochistan. Out of seventy six extra-pulmonary isolates 47 (5%) were from Sindh, 24 (3%) from Punjab and 5 (0.5%) from NWFP.

An age-wise analysis of the *M. tuberculosis* clinical isolates showed that four hundred and forty seven (48%) isolates were from patients aged 15–30 years (214 males, 233 females), 246 isolates (26%) were from patient's aged 31–45 years (141 males, 105 females), while 233 isolates were from the 45+ years (25%) age group (154 males, 79 females).

Spoligotyping of M. tuberculosis isolates

All *M. tuberculosis* strains were spoligotyped and their data analyzed using UPGMA calculation in the Bionumerics software. Seven hundred and twenty one (78%) isolates grouped into 59 different "shared types" while 205 (22%) isolates had 'orphan' spoligotypes. Central Asian strains 61% (n = 568) were found to be the most predominant genotype. Within the CAS genogroup, the CAS1 strain was predominant while, different levels of similarity were noted between CAS1 and other strains in the study population; 17 strains had 96% similarity, 22 strains had 92%, and 29 strains had 88% similarity to CAS1. A further 124 (13%) isolates were found to be part of the CAS super family of strains.

The shared types identified in this study are presented in Fig 1 in order of descending cluster size. We have previously identified cluster groups Pak 1–10 in our population [32]. An additional 26 clusters, Pak 11–36 not

Class [a]	N (% total) [d]	Rules [b]	Binary [c]	Octal
CAS	**568 (61)**			
ST26 CAS1	411 (44)	Δ4-7,Δ23-34		703777740003771
ST288 CAS2	3 (0.3)	Δ4-10,Δ23-34		700377740003771
ST21 CAS1_KILI	1 (0.1)	Δ4-7,Δ10,Δ20-35		703377400041771
ST25 CAS1_DELHI	15 (2)	Δ4-7,Δ23-34,Δ37-38		703777740003171
ST754 CAS1_DELHI	4 (0.4)	Δ2,Δ4-7,Δ23-34		503777740003771
ST794 CAS1_DELHI	5 (0.5)	Δ4-7,Δ14,Δ23-34		703757740003771
ST381 CAS1_DELHI	1 (0.1)	Δ4-7,Δ23-34,Δ37-39		703777740003071
ST428 CAS	2 (0.2)	Δ4-7,Δ23-34,Δ37		703777740003371
ST486 CAS	7 (0.7)	Δ4-7,Δ23-37		703777740000371
ST357 CAS	15 (2)	Δ4-7,Δ23-36		703777740000771
ST1327CAS	2 (0.2)	Δ4-7,Δ12-13,Δ23-34		703637740003771
ST142 CAS	5 (0.5)	Δ4-7,Δ22-34		703777700003771
ST1093CAS	3 (0.3)	Δ4-7,Δ21-34		703777600003771
ST203 CAS	5 (0.5)	Δ4-7,Δ23-35		703777740001771
Orphan Pak 1	2 (0.2)	Δ4-7,Δ23-36,Δ42-43		703777740000760
Orphan Pak 2	2 (0.2)	Δ4-7,Δ15,Δ23-34		703767740003771
Orphan Pak 3	2 (0.2)	Δ4-7,Δ23-34,Δ39		703777740003671
Orphan Pak 4	2 (0.2)	Δ4-7,Δ19,Δ23-34		703777340003771
Orphan Pak 5	2 (0.2)	Δ4-7,Δ24-34		703777760003771
Orphan Pak 6	2 (0.2)	Δ4-7,Δ12,Δ23-34		703677740003771
Orphan Pak 7	2 (0.2)	Δ4-7,Δ23-34,Δ38-39		703777740003471
Orphan Pak 8	2 (0.2)	Δ4-7,Δ23-36,Δ39		703777740000671
Orphan Pak 9	2 (0.2)	Δ4-7,Δ20-34		703777400003771
Orphan Pak 10	2 (0.2)	Δ4-10,Δ23-34Δ38		700377740003571
Orphan Pak 11	2 (0.2)	Δ4-10,Δ23-36		700377740000771
Orphan Pak 12	2 (0.2)	Δ4-7,Δ15-17,Δ19-34		703761000003771
ST458 U	24 (2.5)	Δ29-34		777777777403771
ST100 MANU1	19 (2)	Δ34		777777777773771
ST54 MANU2	16 (2)	Δ33-34		777777777763771
East African-Indian	**42 (4)**			
ST11 EA13	15 (2)	Δ2-3,Δ23-29,Δ34,Δ37-39		477777777413071
ST126 EA15	4 (0.4)	Δ4-7,Δ23-24,Δ34-37		703777747770371
ST236 EA15	15 (2)	Δ29-32,Δ34		777777777413771
Orphan Pak 15	2 (0.2)	Δ2-3,Δ29-34		477777777403771
Orphan Pak 27	3 (0.3)	Δ4-7,Δ23-24,Δ34		703777747773771
Orphan Pak 33	3 (0.4)	Δ2-3,Δ29-34,Δ37-39		477777777403071
Beijing	**25(3)**			
ST1 Beijing	25 (3)	Δ1-34		000000000003771
Haarlem	**15(2)**			
ST50 Haarlem3	1 (0.1)	Δ31,Δ33-36		777777777720771
ST127 Haarlem4	10 (1.0)	Δ2,Δ29-31,Δ33-36		577777777420771
Orphan Pak 16	2 (0.2)	Δ1,Δ29-32,Δ34,Δ40		577777777420771
Orphan Pak 35	2 (0.2)	Δ29-31,Δ33-34		777777777423771
T	**23 (2)**			
ST53 T1	19 (2)	Δ33-36		777777777760771
ST52 T2	1 (0.1)	Δ33-36,Δ40		777777777760731
ST37 T3	1 (0.1)	Δ13,Δ33-36		777737777760771
Orphan Pak 25	2 (0.2)	Δ2-3,Δ33-34		477777777763771
Latin American-Mediterranean	**2 (0.2)**			
ST64 LAM6	1 (0.1)	Δ21-24,Δ29 Δ33-36		777777307530771
ST42 LAM9	1 (0.1)	Δ21-24, Δ33-36		777777307730771
Other Orphan-Pak [e]	**46 (5)**			
Orphan Pak 13	2 (0.2)	Δ4-7,Δ15-17,Δ19-39		70376100000071
Orphan Pak 14	2 (0.2)	Δ1-3,Δ30-32,Δ34		077777777613071
Orpahn Pak 17	2 (0.2)	Δ2-3,Δ8-13,Δ15-16,Δ30-32,Δ34,Δ37-39		236011777613071
Orpahn Pak 18	2 (0.2)	Δ4-8,Δ15,Δ31-41		701767777700011
Orphan Pak 19	2 (0.2)	Δ10,Δ33-37		777377777760371
Orphan Pak 20	2 (0.2)	Δ33-36,Δ39-40		777777777760631
Orphan Pak 21	2 (0.2)	Δ2,Δ19,Δ29-31,Δ33-36		577777377420771
Orphan Pak 22	2 (0.2)	Δ29-31,Δ33-35		777777777421771
Orphan Pak 23	2 (0.2)	Δ26-31,Δ33-37		777777777020771
Orphan Pak 24	2 (0.2)	Δ29-31,Δ33-36		777777777420771
Orphan Pak 26	2 (0.2)	Δ7,Δ21-24,Δ29,Δ33-36		773777607560771
Orphan Pak 28	2 (0.2)	Δ4-7,Δ9-10,Δ23-34		702377740003771
Orpahn Pak 29	2 (0.2)	Δ4-9,Δ23-35		700777600001771
Orphan Pak 30	2(0.2)	Δ29-31,Δ34,Δ37-39		777777777413071
Orphan Pak 31	3 (0.3)	Δ4-7,Δ21,Δ23-34		703777640003771
Orphan Pak 32	5 (0.4)	Δ4-7,Δ15-16, 19-37,Δ40		703763000000331
OrphanPak 34	5 (0.4)	Δ4-7,Δ14-17,Δ19-36,Δ42-43		403741000000760
Orphan Pak 36	5 (0.5)	Δ1-4,Δ12-13,Δ33-36		037637777760771

Figure 1
Spoligotypes shared by Pakistani *M. tuberculosis* isolates. [a] Genotypes as identified in SpolDB4.0 including related Orphan-Pak clusters. [b] Rules as defined by absence of spacers (3). [c] Filled boxes represent positive hybridization while empty boxes represent absence of spacers. [d] Number of isolates. [e] Orphan Pak clusters not identified as genotype of any phylogenetic lineage.

defined within SpolDB4.0 are described here. Orphan Spoligotypes Pak 1–12 displayed 92–98% similarity with CAS1 and its sub-families and were classified as CAS genotype strains. Manu strains characterized by absence of spacer 33–34 have been included in CAS genogroup [10]. These are presumed to be the probable ancestor of CAS strain types. Pak15, Pak27 and Pak33 showed 88% similarity to EAI strains and were classified with East African-Indian genotype. Whereas, Pak16 and Pak35 had 82% similarity to Haarlem genotype strains.

An overall analysis of strain distribution showed that 3/25 (12%) of Beijing strains were isolated from the under 15 years age group as opposed to 8/544 (1.5%) of CAS strains, 3/205 (1.5%) orphan strains and 12/152 (7.9%) of the other shared types (Fig 2). Our data suggests a significant association of Beijing strains with the younger age group (P value = 0.03, Chi square test).

M. tuberculosis genotypes distribution from the four provinces of the country is illustrated in (Fig 3). The largest number of strains was from Sindh followed by Punjab, NWFP and Balochistan. Central Asian strain 1 (CAS1), CAS sub-families (CAS_DEHLI, CAS, U, MANU1 and MANU2) and Beijing strains were prevalent across the country, while EAI strains, T1 family, LAM and Haarlem strains were also present in the different provinces. A comparison of spoligotype distributions between the four regions indicated that the occurrence of CAS1 clade was significantly higher in the Punjab (P < 0.01, Pearsons Chi-square test) as compared with Sindh, NWFP and Balochistan.

Deletion analysis
The TbD1 region has been identified as an evolutionary marker in *M. tuberculosis* strain lineages [41]. Total 36 *M.*

tuberculosis strains were tested for the TbD1 region, 2 strains each from CAS1 and its sub-families (ST25, ST428, ST754, ST794, ST1327, ST288, ST142, ST1093, ST203, ST100, ST54, ST458, and ST21) and in Beijing, Haarlem, EAI3 and EAI5 genotypes. TbD1 region was found to be deleted in all of the above genotypes studied except for the EAI lineages, confirming that the predominant CAS1 strains in this population belongs to a "modern" TB lineage while that EAI strains harboring the TbD1 region and are likely to be "ancestral strains"[43].

Pulmonary versus extra pulmonary isolates
We next determined genotypes of *M. tuberculosis* strains as identified from either pulmonary [n = 850; Sputum (n = 722) and Bronchial lavage (n = 128)] and extra pulmonary sources [n = 76; Pleural fluid (n = 30), Pus (n = 26), Urine (n = 8) and Tissue (n = 12)]. Table 1 lists these in sequence of genotypes given in Fig 1.

There was no statistically significant difference in the distribution of shared spoligotypes i.e in CAS1, Beijing, EAI, Haarlem, T, LAM and other Pak clusters between their pulmonary and extra pulmonary sources. However, 33% of the isolates from extra pulmonary sources were orphan types, compared with 21% of the pulmonary isolates and this difference was statistically significant (Pearsons Chi-square test P = 0.018; 95% CI, 2.0 to 23.5).

Drug resistance patterns
We determined the susceptibility pattern of strains to investigate a relationship between *M. tuberculosis* strain prevalence and drug resistance. Drug resistance patterns of Beijing, LAM, T1, EAI, Haarlem, Orphan Pak, and Orphan types were compared with resistance patterns of CAS genotype (see Table 2 panel-1 In-Bold). Resistance to Streptomycin, Rifampicin, Ethambutol, Isoniazid, Pyrazinamide
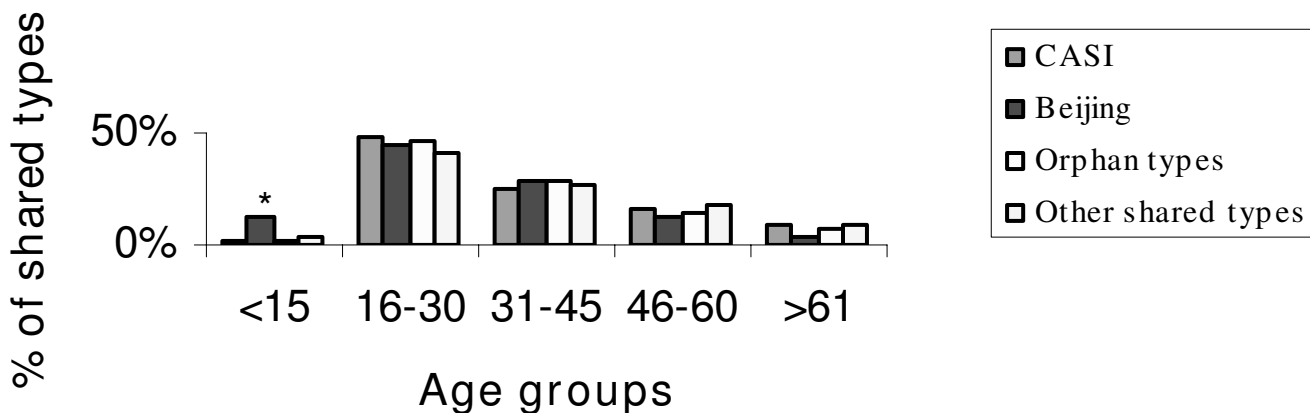
Figure 2
Distribution of *M. tuberculosis* shared types by age. The graph illustrates number of clustered spoligotypes isolated from the different age groups expressed as a percentage of total isolates in that cluster (% of clusters). * Beijing strains showed significant association with < 15 year age group (P = value 0.03 Pearson's Chi Square test).
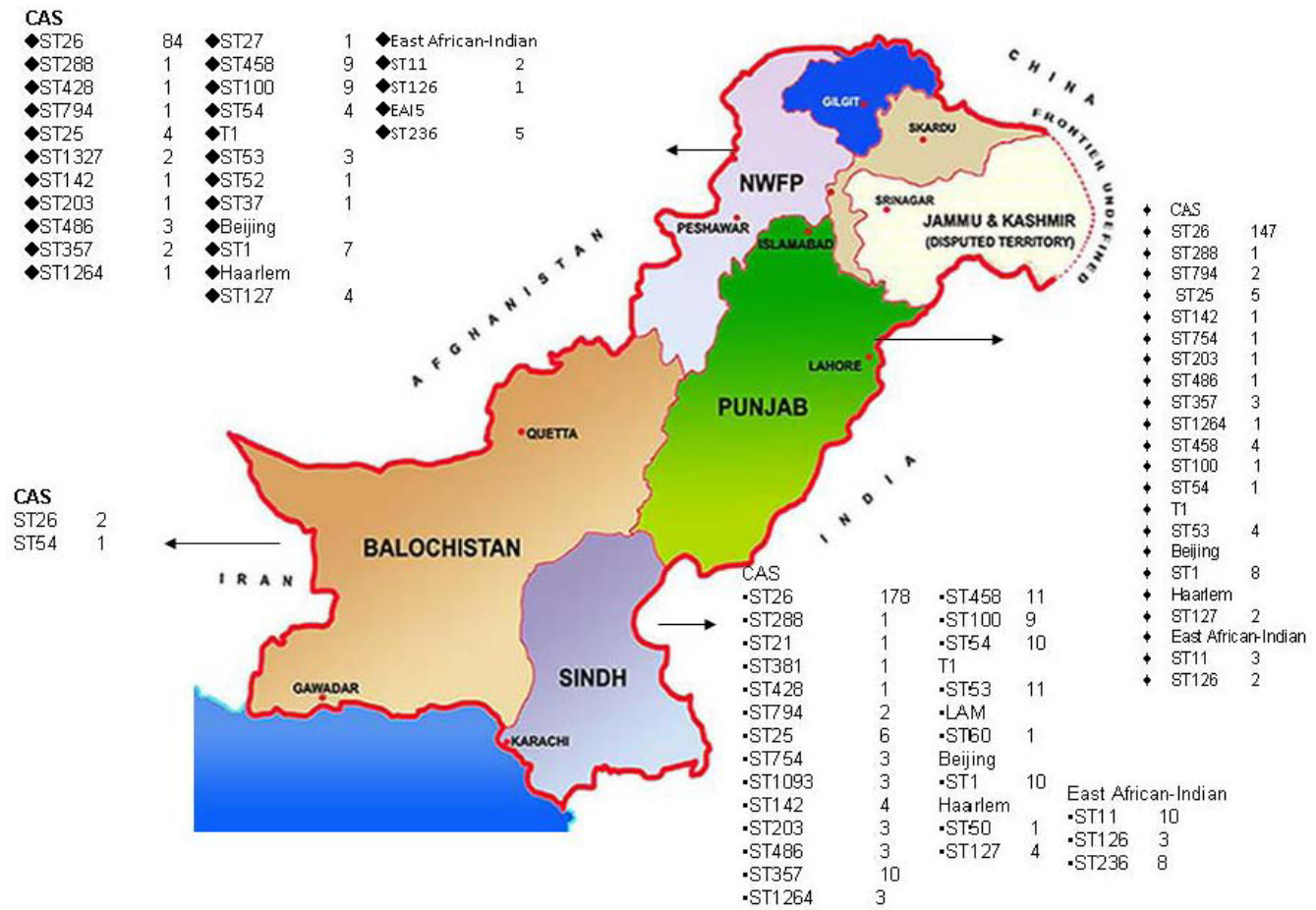
Figure 3
Geographical distribution of *M. tuberculosis* shared types in Pakistan. The map describes the geographical distribution of different genotypes in the Sindh, Punjab, NWFP and Baluchistan provinces. The prevalence of CAS genotype and predominant clusters in each province is depicted. Clades: Beijing ;Central Asian strains CAS; Central Asian sub-families; T1 ill-defined T family; Haarlem H strains; East African Indian strain EAI strains.

was determined in CAS genotype including sub-groups of CAS family and related Pak-clusters. Of the 926 isolates included in this study, 405 (44%) were sensitive to all 5 first line agents tested and 402 (43%) were MDR.

The association between *M. tuberculosis* genotypes and first line drug resistance was assessed by logistic model in comparison with CAS genotype, a significantly higher resistance was observed amongst 'Orphan' isolates and to streptomycin (P value = 0.001, OR; 2.35, CI; 1.68–3.29) and ethambutol (P value = 0.001, OR; 2.15, CI; 1.57–2.93) while, Beijing strains showed higher resistance to streptomycin (P value = 0.041, OR; 2.6; CI; 1.04–6.52).

Analysis of cluster types further showed that Beijing strains (P value = 0.01, Pearsons Chi-square test), and EAI (P value = 0.001, Pearsons Chi-square test) were associated with MDR. The MDR rate in the predominant CAS

genotype was not found to be statistically significant (P value = 0.36, Pearsons Chi-square test).

Discussion
This study presents novel information regarding genotypic diversity and drug resistance of *M. tuberculosis* strains in Pakistan. We found that 411 (44%) of 926 study isolates to be CAS1 or ST26 strains [25], confirming previous reports of 39% prevalence of CAS1 amongst *M. tuberculosis* strains [32]. The results further showed that in addition to ST26 the majority of spoligotypes belonged to the CAS genotype. CAS1 has been also identified by recent studies as a predominant strain in Delhi[25] and Mumbai[44]. Prevalence of Beijing strains in our study at 3% (n = 25) compares well with data from Delhi, where 8% of 105 isolates are reported to be of Beijing family[25]. We further identified 36 clusters (Pak clusters) not identical with any of the STs described within SpolDB4.0.

Table 1: *M. tuberculosis* shared typed distribution amongst pulmonary and extra-pulmonary disease isolates.

Spoligotype *	Pulmonary n(%)	Extra-pulmonary n(%)
CAS	531(62)	37(49)
East African-Indian	37(4)	5(6)
Beijing	24(3)	1(1)
Haarlem	15(2)	0
T	20(2)	3(4)
LAM	2(0.2)	0
Other Orphan-Pak types[a]	41(5)	5(6)
Orphan types[b]	180(21)	25(33)**
Total	**850**	**76**

* Strain genotypes as listed in Fig 1.
[a]Include clustered Orphan-Pak types (Pak 13–14, 17–24, 26, 28–32, 34 and 36)
[b]Include unclustered Orphan types
**Denotes significant difference between pulmonary and extra-pulmonary sources, (p < 0.05, Chi x^2 test)

Comparison of pulmonary and extra pulmonary sources showed a significant association of 'Orphan' spoligotypes with extra pulmonary disease. This difference between pulmonary and extrapulmonary spoligotypes may be attributed to the greater ease with which pulmonary strains may be transmitted via aerosol routes, leading to greater transmission and strain clustering.

The global population structure of *M. tuberculosis* is reportedly defined by six phylogeographical lineages, each associated with specific human populations http://www.pnas.org/cgi/doi/10.1073/pnas.052548299[29,45]. It has been suggested that particular lineages of *M. tuberculosis* might be adapted to specific human populations and maladapted to others[45]. Strain differences in different geographical regions may be linked to different ethnic subpopulations in these regions, and their migration histories [46].

The CAS strains have been shown to be predominant [30,47] in this region, the distribution of CAS1 strains was greater in the province of Punjab (P < 0.01, Pearsons Chi-square test) as compared with Sindh, NWFP, and Balochistan. This could be due to the shared border and similar population mix between Punjab now in Pakistan (West Punjab prior to the 1947 partition of India), and East Punjab, (now Punjab, North India).

Although not statistically significant, East African-Indian strains were more prevalent in Sindh. The fact that 4% of our isolates were East African-Indian is comparable to an earlier study from Delhi, reporting 8% EAI strains in their population [25]. Both these figures are in contrast to a study from Southern India indicating an 80% prevalence of TbD1+/EAI isolates amongst their samples [43]. It has been hypothesized that lineages that are rare in specific human population are not adapted to spreading within these populations and show a significantly lower case-rate

ratio. This hypothesis is supported by a recent report[30] suggesting that TB in India is essentially caused by historic clones of tubercle bacilli which circulate with geographic predilection.

Deletion analysis demonstrated the absence of TbD1 region in CAS1 and its sub-families. A recent study [43] also suggests that CAS1 family evolved as a result of an evolutionary event causing TbD1 deletion from a common ancestor[18,41]. The study further suggests that strains within CAS genotype are related through minor genetic changes. The similarity seen in our study between Pak clusters 1–12 and CAS genotype suggests an evolution from a common ancestor belonging to the same phylogenetic lineage CAS.

Haarlem4 and T1 were also identified amongst our isolates. It is reported that more than 60% of ST127 (H4, modern type) are localized in Iran and Russia. Their presence in our population suggests the role of traditional migratory routes from central Asia in the history of tuberculosis. In addition to the predominant groups, we also found clusters of rare/localized shared types listed in SpolDB4.0 that have previously been found only in North America, Australia and Europe (CAS1_KILI) as well as those found in neighboring Iran and India (LAM 9, LAM 6).

In the absence of national surveillance for drug resistance in Pakistan, the majority of published reports rely upon laboratory data based on passive specimen collection. There is thus little information about the burden of MDR at a community level. WHO estimates suggest an MDR rate of under 3% amongst new cases and 20–40% amongst previously treated TB cases[48]. A recent report of 1.8% MDR-TB, in untreated cases in Pakistan corroborates the WHO estimates [49]. Since the samples in our study were not collected through active case finding, the

Table 2: Drug resistance patterns of predominant *M. tuberculosis* shared types as compared with genotype of CAS lineage.

clusters**	Sensitive n(%)a	Drug resistant isolates (n) as a percentage of total isolates in the clades (%)					
		Rifampicin	Isoniazid	Streptomycin	Ethambutol	Pyrazinamide	Total
CAS	245 (45)	252 (46)	258 (47)	141 (26)	210 (39)	297 (55)	568
EAI	17 (35)	30 (61)	32 (65)	17 (35)	24 (49)	29 (59)	42
P value		0.08	0.06	0.12	0.13	0.28	
Odds ratio		1.71	1.78	1.73	1.64	1.4	
95% CI		0.92–3.19	0.96–3.30	0.86–3.51	0.86–3.14	0.75–2.62	
Beijing	8 (32)	18 (72)	17 (68)	12 (48)	15 (60)	16 (64)	25
P value		0.071	0.1	**0.041***	0.08	0.25	
Odds ratio		2.18	2.01	2.6	2.18	1.65	
95% CI		0.93–5.12	0.85–4.76	1.04–6.52	0.90–5.26	0.69–3.92	
Haarlem	9 (60)	4 (27)	6 (40)	Not resistant	4 (27)	5 (33)	15
P value		0.16	0.25		0.67	0.16	
Odds ratio		0.43	0.52		0.772	0.45	
95% CI		0.13–1.42	0.17–1.59		0.23–2.55	0.15–1.38	
T	6 (30)	8 (40)	9 (45)	5 (25)	5 (25)	7 (35)	23
P value		0.63	0.5	0.54	0.96	0.94	
Odds ratio		1.29	1.42	1.44	0.97	0.96	
95% CI		0.44–3.79	0.5–4.0	0.43–4.8	0.29–3.23	0.31–3.90	
LAM	1 (50)	1 (50)	1 (50)	1 (50)	1 (50)	1 (50)	2
P value		0.98	0.97	0.69	0.91	0.89	
Odds ratio		0.97	0.95	1.73	1.61	0.82	
95% CI		0.06–15.03	0.05–15.26	0.10–27.99	0.07–18.76	0.05–13.25	
Other Orphan-Pakb	20 (38)	19 (36)	21 (40)	9 (17)	19 (36)	13 (24)	46
P value		0.93	0.99	0.55	0.75	0.08	
Odds ratio		0,97	0.99	0.78	1.1	0.53	
95% CI		0.51–1.85	0.52–1.88	0.34–1.76	0.57–2.13	0.26–1.1	
Orphan types	96 (47)	106 (52)	137 (67)	130 (63)	177 (86)	129 (63)	205
P value		0.67	0.058	**0.001***	**0.001***	0.52	
Odds ratio		1.03	1.35	2.35	2.15	1.1	
95% CI		0.77–1.48	0.99–1.85	1.68–3.29	1.57–2.93	0.81–1.1	

aSensitive to first line drugs (Rifampicin, Isoniazid, Streptomycin, Ethambutol and Pyrazinamide)
bIncludes a total of 18 clusters
**Strain genotypes as listed in Fig. 1
* Significant difference compared to CAS genotypes

high MDR rate noted in our specimens is likely to reflect the large number of previously treated patients included in this study. This however, cannot be confirmed due to absence of prior treatment history for our patients.

In agreement with an earlier report [32] CAS1 and related CAS sub-family spoligotypes in this study showed no correlation with MDR-TB. However, a higher relative risk of MDR amongst Beijing strains was noted. Association between Beijing strains and MDR varies worldwide [50,51], A recent study from Mumbai, India reported a

higher frequency of Beijing strains (35%) amongst MDR isolates [44]. The distinctive feature of Beijing MDR-TB outbreak is accelerated transmission as compared with other MDR-TB outbreaks[52]. Further, East African-Indian strain was also found to be associated with MDR supported by a study from Iran [53]. Recent studies moreover suggests that MDR-TB strains may be responsible for emergence of XDR-TB cases [53,54].

The higher occurrence of streptomycin resistance in Beijing strains as compared to CAS genotype has importance

in that particular lineage of *M. tuberculosis* might harbor polymorphisms which make them resistant to certain anti-tuberculosis drugs.

In our study population, association of Beijing strains with a younger age group (<15 years) is concerning. It suggests that Beijing strains may be an emerging strain type and is likely to increasingly contribute to the burden of drug resistant tuberculosis in this region.

Conclusion

The identification of a dominant spoligotype, CAS, similar to previously identified isolates in India and Bangladesh illustrates an important trend in the *M. tuberculosis* infection pattern in the South Asian region. All predominant clusters apart from CAS1 strains were equally distributed in the country suggest a continual transmission of strains. This data presents a comprehensive evaluation of the strain-to-strain variability in *M. tuberculosis*, important in phenotypic consequences, also these phylogeographical strain variation may affect the development of new diagnostic tools, drugs, and vaccines for treatment in the endemic region.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MT conducted the experimental work and prepared the manuscript. ZH supervised the laboratory work and study design. AK and AA provided technical help with spoligotyping. RS provided statistical data analysis support. SG provided advice regarding MTB strain typing. RH planned the idea and supervised the study. All authors read and approved the final manuscript.

Acknowledgements

References

1.   Dye C, Scheele SDP, Pathania V, Raviglione MC: **Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. World Health Organization Global Surveillance and Monitoring project.** *JAMA* 1999, **282:**677-686.
2.   Bloom BR, Murray CJ: **Tuberculosis: commentary on a reemergent killer.** *Science* 1992, **257(5073):**1055-1064.
3.   WHO: **WHO report. Global Tuberculosis control: Surveillance, Planning, Financing.** 2007 [http://www.who.int/tb/publications/global_report].

4.   Ahmad K: **Stop TB partnership to focus on Afghanistan and Pakistan.** *Lancet* 2001, **358(9291):**1434.
5.   WHO: **Guidlines for the programmatic management of drug-resistant tuberculosis.** 2007.
6.   Butt T, Ahmad RN, Kazmi SY, Rafi N: **Multi-drug resistant tuberculosis in Northern Pakistan.** *J Pak Med Assoc* 2004, **54(9):**469-472.
7.   Karamat KA, Rafi S, Abbasi SA: **Drug resistance in Mycobacterium tuberculosis: a four years experience.** *J Pak Med Assoc* 1999, **49(11):**262-265.
8.   Irfan S, Hassan Q, Hasan R: **Assessment of resistance in multi drug resistant tuberculosis patients.** *J Pak Med Assoc* 2006, **56(9):**397-400.
9.   Sola C, Filliol I, Gutierrez MC, Mokrousov I, Vincent V, Rastogi N: **Spoligotype database of Mycobacterium tuberculosis: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives.** *Emerg Infect Dis* 2001, **7(3):**390-396.
10. Filliol I, Driscoll JR, van Soolingen D, Kreiswirth BN, Kremer K, Valetudie G, Dang DA, Barlow R, Banerjee D, Bifani PJ, *et al.*: **Snapshot of moving and expanding clones of Mycobacterium tuberculosis and their global distribution assessed by spoligotyping in an international study.** *J Clin Microbiol* 2003, **41(5):**1963-1970.
11. Banu S, Gordon SV, Palmer S, Islam MR, Ahmed S, Alam KM, Cole ST, Brosch R: **Genotypic analysis of Mycobacterium tuberculosis in Bangladesh and prevalence of the Beijing strain.** *J Clin Microbiol* 2004, **42(2):**674-682.
12. Sun YJ, Bellamy R, Lee AS, Ng ST, Ravindran S, Wong SY, Locht C, Supply P, Paton NI: **Use of mycobacterial interspersed repetitive unit-variable-number tandem repeat typing to examine genetic diversity of Mycobacterium tuberculosis in Singapore.** *J Clin Microbiol* 2004, **42(5):**1986-1993.
13. Zozio T, Allix C, Gunal S, Saribas Z, Alp A, Durmaz R, Fauville-Dufaux M, Rastogi N, Sola C: **Genotyping of Mycobacterium tuberculosis clinical isolates in two cities of Turkey: description of a new family of genotypes that is phylogeographically specific for Asia Minor.** *BMC microbiology* 2005, **5:**44.
14. Kato-Maeda M, Bifani PJ, Kreiswirth BN, Small PM: **The nature and consequence of genetic variability within Mycobacterium tuberculosis.** *The Journal of clinical investigation* 2001, **107(5):**533-537.
15. Toungoussova OS, Sandven P, Mariandyshev AO, Nizovtseva NI, Bjune G, Caugant DA: **Spread of drug-resistant Mycobacterium tuberculosis strains of the Beijing genotype in the Archangel Oblast, Russia.** *J Clin Microbiol* 2002, **40(6):**1930-1937.
16. Kubin M, Havelkova M, Hyncicova I, Svecova Z, Kaustova J, Kremer K, van Soolingen D: **A multidrug-resistant tuberculosis microepidemic caused by genetically closely related Mycobacterium tuberculosis strains.** *J Clin Microbiol* 1999, **37(8):**2715-2716.
17. Lopez B, Aguilar D, Orozco H, Burger M, Espitia C, Ritacco V, Barrera L, Kremer K, Hernandez-Pando R, Huygen K, *et al.*: **A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes.** *Clin Exp Immunol* 2003, **133(1):**30-37.
18. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, Musser JM: **Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination.** *Proc Natl Acad Sci USA* 1997, **94(18):**9869-9874.
19. Kemp EB, Belshe RB, Hoft DF: **Immune responses stimulated by percutaneous and intradermal bacille Calmette-Guerin.** *J Infect Dis* 1996, **174(1):**113-119.
20. Behr MA, Small PM: **A historical and molecular phylogeny of BCG strains.** *Vaccine* 1999, **17(7–8):**915-922.
21. Andersen P, Doherty TM: **The success and failure of BCG – implications for a novel tuberculosis vaccine.** *Nat Rev Microbiol* 2005, **3(8):**656-662.
22. Davids V, Hanekom WA, Mansoor N, Gamieldien H, Gelderbloem SJ, Hawkridge A, Hussey GD, Hughes EJ, Soler J, Murray RA, *et al.*: **The effect of bacille Calmette-Guerin vaccine strain and route of administration on induced immune responses in vaccinated infants.** *J Infect Dis* 2006, **193(4):**531-536.
23. Goyal M, Saunders NA, van Embden JD, Young DB, Shaw RJ: **Differentiation of Mycobacterium tuberculosis isolates by spoligotyping and IS6110 restriction fragment length polymorphism.** *J Clin Microbiol* 1997, **35(3):**647-651.

24. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, *et al*.: **Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology.** *J Clin Microbiol* 1997, **35(4):**907-914.
25. Singh UB, Suresh N, Bhanu NV, Arora J, Pant H, Sinha S, Aggarwal RC, Singh S, Pande JN, Sola C, *et al*.: **Predominant tuberculosis spoligotypes, Delhi, India.** *Emerg Infect Dis* 2004, **10(6):**1138-1142.
26. Doroudchi M, Kremer K, Basiri EA, Kadivar MR, van Soolingen D, Ghaderi AA: **IS6110-RFLP and spoligotyping of Mycobacterium tuberculosis isolates in Iran.** *Scandinavian journal of infectious diseases* 2000, **32(6):**663-668.
27. Qian L, Van Embden JD, Zanden AG Van Der, Weltevreden EF, Duanmu H, Douglas JT: **Retrospective analysis of the Beijing family of Mycobacterium tuberculosis in preserved lung tissues.** *J Clin Microbiol* 1999, **37(2):**471-474.
28. Bhanu NV, van Soolingen D, van Embden JD, Dar L, Pandey RM, Seth P: **Predominace of a novel Mycobacterium tuberculosis genotype in the Delhi region of India.** *Tuberculosis (Edinb)* 2002, **82(2–3):**105-112.
29. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, Allix C, Aristimuno L, Arora J, Baumanis V, *et al*.: **Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology.** *BMC microbiology* 2006, **6:**23.
30. Singh UB, Arora J, Suresh N, Pant H, Rana T, Sola C, Rastogi N, Pande JN: **Genetic biodiversity of Mycobacterium tuberculosis isolates from patients with pulmonary tuberculosis in India.** *Infect Genet Evol* 2007, **7(4):**441-448.
31. Velayati AA, Farnia P, Mirsaeidi M, Reza Masjedi M: **The most prevalent Mycobacterium tuberculosis superfamilies among Iranian and Afghan TB cases.** *Scandinavian journal of infectious diseases* 2006, **38(6–7):**463-468.
32. Hasan Z, Tanveer M, Kanji A, Hasan Q, Ghebremichael S, Hasan R: **Spoligotyping of Mycobacterium tuberculosis isolates from Pakistan reveals predominance of Central Asian Strain 1 and Beijing isolates.** *J Clin Microbiol* 2006, **44(5):**1763-1768.
33. Mardassi H, Namouchi A, Haltiti R, Zarrouk M, Mhenni B, Karboul A, Khabouchi N, Gey van Pittius NC, Streicher EM, Rauzier J, *et al*.: **Tuberculosis due to resistant Haarlem strain, Tunisia.** *Emerg Infect Dis* 2005, **11(6):**957-961.
34. Glynn JR, Whiteley J, Bifani PJ, Kremer K, van Soolingen D: **Worldwide occurrence of Beijing/W strains of Mycobacterium tuberculosis: a systematic review.** *Emerg Infect Dis* 2002, **8(8):**843-849.
35. Nolte FS, Metchock B: **Mycobacterium.** *Manual of clinical microbiology* 6th edition. 1995:400-437.
36. Wayne LGalK: **Preperation of tuberculosis susceptibility testing mediums by impregnated discs.** *AmJPublic Health* 1966, **45:**769-771.
37. Villanova P: **Antimycobacterial susceptibility testing for Mycobacterium tuberculosis.** *National Committee for Clinical Laboratory Standards15M24T[15]* 1995.
38. Isenberg HD: **Clinical microbiology procedure handbook.** *Volume 2.* 2nd edition. ASM Press; 2004:7821-7823.
39. Honore-Bouakline S, Vincensini JP, Giacuzzo V, Lagrange PH, Herrmann JL: **Rapid diagnosis of extrapulmonary tuberculosis by PCR: impact of sample preparation and DNA extraction.** *J Clin Microbiol* 2003, **41(6):**2323-2329.
40. Baker L, Brown T, Maiden MC, Drobniewski F: **Silent nucleotide polymorphisms and a phylogeny for Mycobacterium tuberculosis.** *Emerg Infect Dis* 2004, **10(9):**1568-1577.
41. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, *et al*.: **A new evolutionary scenario for the Mycobacterium tuberculosis complex.** *Proc Natl Acad Sci USA* 2002, **99(6):**3684-3689.
42. Flores L, Van T, Narayanan S, DeRiemer K, Kato-Maeda M, Gagneux S: **Large sequence polymorphisms classify Mycobacterium tuberculosis strains with ancestral spoligotyping patterns.** *J Clin Microbiol* 2007, **45(10):**3393-3395.
43. Gutierrez MC, Ahmed N, Willery E, Narayanan S, Hasnain SE, Chauhan DS, Katoch VM, Vincent V, Locht C, Supply P: **Predominance of ancestral lineages of Mycobacterium tuberculosis in India.** *Emerg Infect Dis* 2006, **12(9):**1367-1374.
44. Almeida D, Rodrigues C, Ashavaid TF, Lalvani A, Udwadia ZF, Mehta A: **High incidence of the Beijing genotype among multidrug-resistant isolates of Mycobacterium tuberculosis in a tertiary care center in Mumbai, India.** *Clin Infect Dis* 2005, **40(6):**881-886.
45. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, *et al*.: **Variable host-pathogen compatibility in Mycobacterium tuberculosis.** *Proc Natl Acad Sci USA* 2006, **103(8):**2869-2873.
46. Singh M, Balamurugan A, Katoch K, Sharma SK, Mehra NK: **Immunogenetics of mycobacterial infections in the North Indian population.** *Tissue Antigens* 2007, **69(Suppl 1):**228-230.
47. Gagneux S, Small PM: **Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development.** *Lancet Infect Dis* 2007, **7(5):**328-337.
48. WHO: **Anti-tuberculosis drug resistance in the world, forth global report.** World Health Organization Geneva, Switzerland; 2008.
49. Javaid A, Hasan R, Zafar A, Ghafoor A, Pathan AJ, Rab A, Sadiq A, Akram CM, Burki I, Shah K, *et al*.: **Prevalence of primary multidrug resistance to anti-tuberculosis drugs in Pakistan.** *Int J Tuberc Lung Dis* 2008, **12(3):**326-331.
50. Anh DD, Borgdorff MW, Van LN, Lan NT, van Gorkom T, Kremer K, van Soolingen D: **Mycobacterium tuberculosis Beijing genotype emerging in Vietnam.** *Emerg Infect Dis* 2000, **6:**302-305.
51. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P, van Embden JD: **Predominance of a single genotype of Mycobacterium tuberculosis in countries of east Asia.** *J Clin Microbiol* 1995, **33(12):**3234-3238.
52. Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN: **Global dissemination of the Mycobacterium tuberculosis W-Beijing family strains.** *Trends Microbiol* 2002, **10(1):**45-52.
53. Masjedi MR, Farnia P, Sorooch S, Pooramiri MV, Mansoori SD, Zarifi AZ, Akbarvelayati A, Hoffner S: **Extensively drug-resistant tuberculosis: 2 years of surveillance in Iran.** *Clin Infect Dis* 2006, **43(7):**841-847.
54. Hasan Z, Mahmood A, Zafar S, Khan AA, Hussain R: **Leprosy patients with lepromatous disease have an up-regulated IL-8 response that is unlinked to TNF-alpha responses.** *Int J Lepr Other Mycobact Dis* 2004, **72(1):**35-44.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2334/8/171/prepub