Research

Microevolution of extensively drug-resistant tuberculosis in Russia

Nicola Casali,[1] Vladyslav Nikolayevskyy,[1] Yanina Balabanova,[1] Olga Ignatyeva,[2]
Irina Kontsevaya,[2] Simon R. Harris,[3] Stephen D. Bentley,[3] Julian Parkhill,[3]
Sergey Nejentsev,[4] Sven E. Hoffner,[5] Rolf D. Horstmann,[6] Timothy Brown,[1]
and Francis Drobniewski[1,7]

[1]National Mycobacterium Reference Laboratory, Blizard Institute, Queen Mary, University of London, London E1 2AT, United Kingdom;
[2]Samara Oblast Tuberculosis Dispensary, 443068 Samara, Russian Federation; [3]The Wellcome Trust Sanger Institute, Wellcome Trust
Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; [4]Department of Medicine, University of Cambridge, Cambridge
CB2 2QQ, United Kingdom; [5]Department of Bacteriology, Swedish Institute for Infectious Disease Control, 17182 Solna, Sweden;
[6]Department of Molecular Medicine, Bernhard Nocht Institute for Tropical Medicine, 20359 Hamburg, Germany

Extensively drug-resistant (XDR) tuberculosis (TB), which is resistant to both first- and second-line antibiotics, is an escalating problem, particularly in the Russian Federation. Molecular fingerprinting of 2348 *Mycobacterium tuberculosis* isolates collected in Samara Oblast, Russia, revealed that 72% belonged to the Beijing lineage, a genotype associated with enhanced acquisition of drug resistance and increased virulence. Whole-genome sequencing of 34 Samaran isolates, plus 25 isolates representing global *M. tuberculosis* complex diversity, revealed that Beijing isolates originating in Eastern Europe formed a monophyletic group. Homoplasic polymorphisms within this clade were almost invariably associated with antibiotic resistance, indicating that the evolution of this population is primarily driven by drug therapy. Resistance genotypes showed a strong correlation with drug susceptibility phenotypes. A novel homoplasic mutation in *rpoC*, found only in isolates carrying a common *rpoB* rifampicin-resistance mutation, may play a role in fitness compensation. Most multidrug-resistant (MDR) isolates also had mutations in the promoter of a virulence gene, *eis*, which increase its expression and confer kanamycin resistance. Kanamycin therapy may thus select for mutants with increased virulence, helping preserve bacterial fitness and promoting transmission of drug-resistant TB strains. The East European clade was dominated by two MDR clusters, each disseminated across Samara. Polymorphisms conferring fluoroquinolone resistance were independently acquired multiple times within each cluster, indicating that XDR TB is currently not widely transmitted.

[Supplemental material is available for this article.]

Multidrug-resistant (MDR) tuberculosis (TB), which is resistant to at least isoniazid and rifampicin, is a growing threat to national tuberculosis programs worldwide. MDR TB patients are more difficult and costly to treat and remain infectious for longer, and rates of default, failure and death are higher. Extensive resistance (XDR), defined as MDR plus resistance to any fluoroquinolone and at least one of the second-line injectable agents (kanamycin, capreomycin, or amikacin), further decreases chances of treatment success and survival (Drobniewski et al. 2005; Gandhi et al. 2006; Migliori et al. 2008; Shah et al. 2008; Kliiman and Altraja 2009; Wright et al. 2009; Leimane et al. 2010). So far, co-infection of HIV and XDR TB has proved largely fatal (Gandhi et al. 2006).

The problem of drug-resistant TB is particularly acute in countries of the former Soviet Union, where almost half of all cases are resistant to at least one drug; the MDR TB rate exceeds 10% among new cases; and XDR TB can account for up to a quarter of MDR TB cases (Drobniewski et al. 2005; Kliiman and Altraja 2009; Wright et al. 2009; Leimane et al. 2010).

In Samara Oblast (a region of the Russian Federation with a population of 3.3 million), TB incidence and prevalence rates reached 86.8 and 162.5 per 100,000 population in 2009. Extremely high rates of MDR TB (20% among new and 43% among previously treated cases) are coupled with a fast growing epidemic of HIV with the number of HIV-positive cases reaching 40,000 in 2010 (Drobniewski et al. 2005; Samara Oblast Tuberculosis Dispensary 2009; Samara AIDS Center Epidemiological Update 2010 [http://www.noaids.ru/r/11/]). Prolonged hospitalization (minimum, 3 mo) is compulsory for smear/culture-positive patients, and HIV-infected patients stay at the same hospitals with other patients, including drug-resistant cases. Institutional infection control measures are not fully implemented, providing opportunities for MDR/XDR TB super-infection of patients.

On the basis of single nucleotide polymorphism (SNP) analysis, Baker et al. (2004) determined that the global *Mycobacterium tuberculosis* complex (MTBC) population could be divided into five core lineages, here designated Beijing, CAS, EuroAmerican, EAI, and *Mycobacterium bovis* (more typically an animal pathogen). By use of large sequence polymorphisms (LSPs), Gagneux et al. (2006a) confirmed these groupings and defined two *Mycobacterium africanum* lineages: West African-1 and -2. In a previous study, we found that approximately two-thirds of TB cases in Samara were caused by strains belonging to the Beijing lineage (Drobniewski

et al. 2005). There was evidence that Beijing isolates were associated with drug resistance and caused radiologically more advanced disease. Several other studies have also linked the Beijing genotype with greater virulence (for review, see Nicol and Wilkinson 2008).

In this study, we used a whole-genome sequencing approach to examine the microevolution of Beijing strains, collected in Samara, focusing on the acquisition and spread of drug resistance.

Results

Molecular fingerprinting of *M. tuberculosis* strains circulating in Samara

We analyzed the *M. tuberculosis* strain family distribution of 2348 isolates (one per patient) collected in 2008–2010 from across Samara Oblast. In this population, by spacer-oligonucleotide (spoligotype) analysis, 71.8% of strains shared the Beijing genotype. The remainder belonged mostly to the EuroAmerican lineage (22.4%), including the T, Haarlem, and LAM families. CAS and EAI lineages accounted for 0.5% of the analyzed population (Fig. 1A).

All Beijing isolates ($n = 1,685$) were typed using a panel of eight variable number tandem repeat (VNTR) loci, which are highly discriminatory for this lineage (Nikolayevskyy et al. 2006). Clustering based on identity at all eight loci revealed that two VNTR types, containing 341 (20.2%) and 160 (9.5%) isolates, were dominant in the population (Fig. 1B). A further 692 (41.1%) isolates clustered into groups of size 72 to two, while 492 (29.2%) had fingerprints that were unique in the collection.

Selection of MTBC isolates for genome sequencing

Two EuroAmerican and 32 Beijing isolates from Samara were chosen for sequencing. Selection of Beijing isolates, based on VNTR data, covered the breadth of diversity, as well as deep sequencing of common VNTR clones in approximate proportion to their epidemiological abundance (Fig. 1B). One Beijing strain originally isolated from Estonia was also selected. Over 10 years, this strain grew from causing a few cases, to dominate the country's *M. tuberculosis* population (Kruuner et al. 2001). Brown et al. (2010) have demonstrated that global MTBC diversity is reflected in

strains circulating in London. Based on VNTR typing, 23 London isolates were selected to represent maximum variability across the entire MTBC and particularly within the Beijing lineage. Finally, the reference strain, H37Rv, was resequenced.

SNPs and phylogenetic analysis

Sequence reads for each of the 58 MTBC isolates were mapped against the reference genome, *M. tuberculosis* H37Rv. A total of 11,854 polymorphic sites were identified, an average of one every 386 bp (Supplemental Data 1). A phylogeny constructed from these SNPs was congruent with the spoligotype families and supports that previously established through analyses of SNPs and LSPs (Fig. 2; Baker et al. 2004; Gagneux et al. 2006a). The four *M. tuberculosis* lineages are well defined, and isolates from different lineages are separated by about 2000 SNPs. Isolates from within a lineage are separated up to 800 SNPs, with the EuroAmerican strains appearing the most diverse. Eighty-three SNPs shared by all members of the four *M. tuberculosis* lineages discriminated them from the *M. africanum* and *M. bovis* isolates.
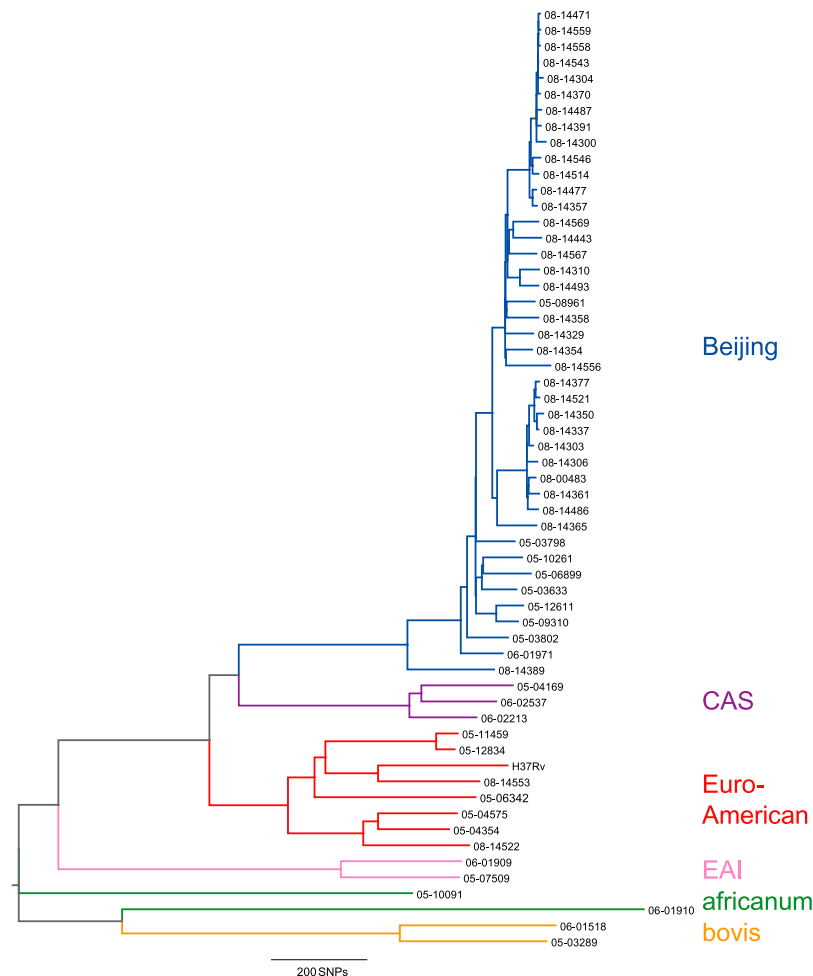
In relation to the diverse London Beijing isolates, with one exception, the Beijing isolates from Samara formed a single clade. The Estonian isolate and one London isolate, from a patient of Eastern European ethnicity, fell into the Samara clade. Comparison of the SNPs we identified, with those determined to define modern versus ancient Beijing subtypes (Schürch et al. 2011), revealed that the exceptional Samara isolate belonged to the ancient subtype, whereas all other Beijing isolates, including those from London, were of the modern subtype.

Structural variation

Three large deletions shared by East European Beijing isolates corresponded to RD105, RD207, and RD181, which define the modern Beijing sublineage (Gagneux et al. 2006a). No large deletions specific to either East European Beijing type, clade A or clade B, were discovered. Unmapped reads were assembled de novo, and resulting contigs were compared with completed *M. tuberculosis* genome sequences. All contig sequences were found in strains belonging to the EuroAmerican lineage, indicating that the evolutionary events detected correspond to deletions within that lineage or specifically H37Rv. The observed lack of horizontal DNA acquisition is consistent with previous reports (Hirsh et al. 2004).
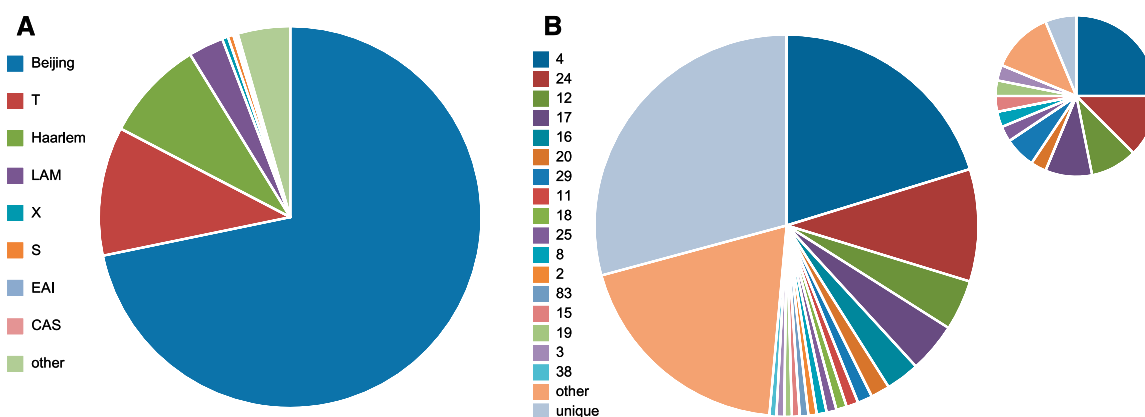
Figure 1. Molecular fingerprinting of Samaran *M. tuberculosis* isolates. (*A*) Prevalence of spoligotype families ($n = 2350$) (Vitol et al. 2006). (*B*) Prevalence of Beijing family VNTR types with cluster size of 10 or more ($n = 1685$). (*Inset*) VNTR types of isolates selected for sequencing ($n = 32$).

Figure 2. Maximum likelihood phylogenetic tree based on 11,854 polymorphic loci. Clades are colored by lineage based on spoligotype analysis. The EuroAmerican lineage comprises T, Haarlem, and LAM spoligotype families.

A, we estimate that ~25% of the Beijing family in Samara belongs to this clade. Clade B isolates shared six of eight VNTR loci and included all the sequenced isolates of VNTR cluster 24, the second-most common type in the population, plus two other clusters. Clade B is thus estimated to represent >11% of circulating Beijing strains. Neither clade showed any association with the geographic origin of the patient within Samara (Supplemental Fig.).

The profound differences in the number of SNPs between and within different VNTR clusters indicates that, in this high incidence region, VNTR typing alone does not accurately reflect true clonal prevalence, significantly underestimating the dominance of the commonest strains.
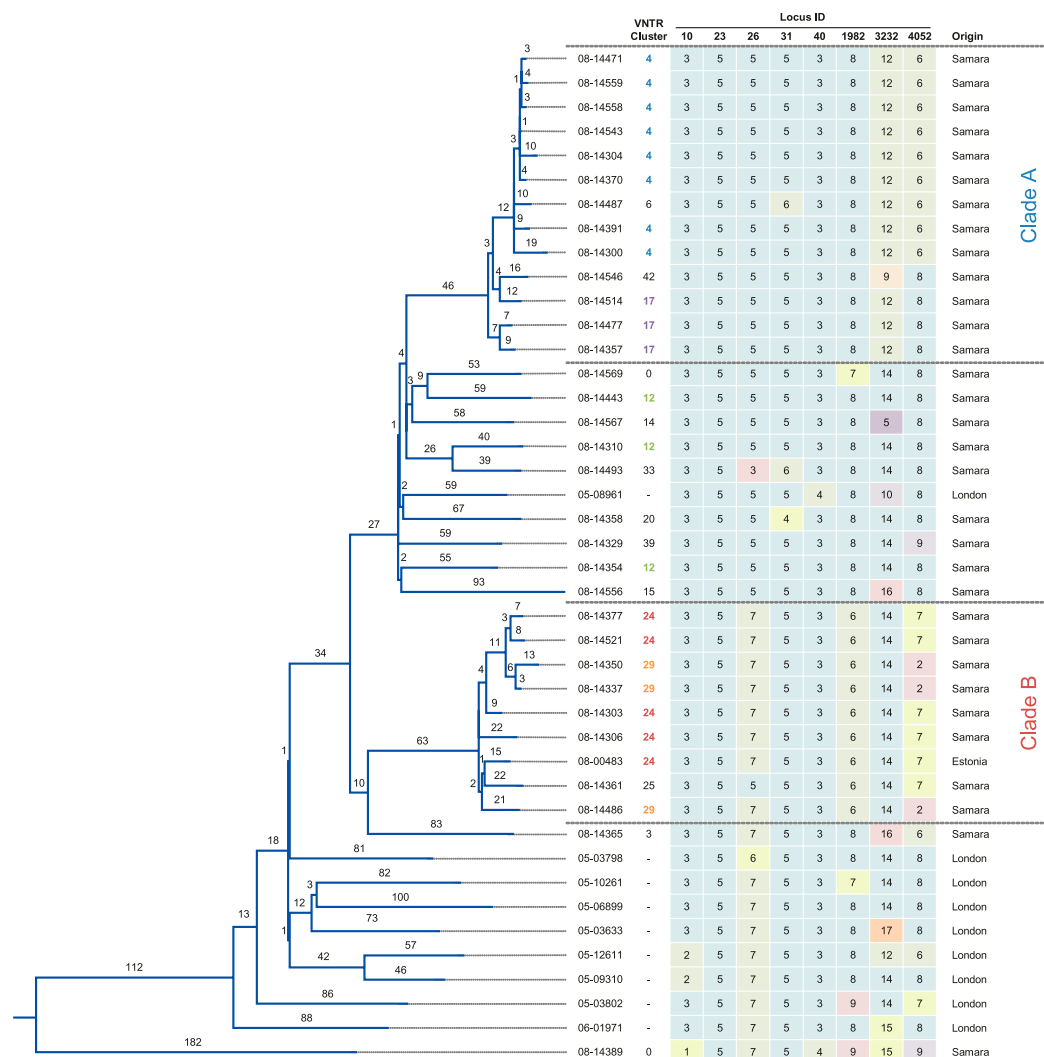
Adaptive selection

M. tuberculosis is an obligate human pathogen, and evolutionary pressure is likely dominantly mediated by the host immune response and antibiotic therapy. Homoplasies, indicating sites of convergent evolution, are a dependable signal of selective pressure. Eighteen homoplasic SNPs were identified within the Beijing lineage (Table 1). SNPs identified in *rpoB*, *katG*, *pncA* promoter, *embB* or its promoter, *rpsL*, *rrs*, and *gyrA* are all common drug resistance mutations. Three homoplasic sites were clustered within the *eis* promoter, lending substantial support to the reported involvement of this region in conferring kanamycin resistance (Zaunbrecher et al. 2009). The final homoplasic SNP was located within *rpoC*, which encodes the RNA polymerase (RNAP) β′ subunit. It was found only in isolates that carried the rifampicin-resistance–conferring mutation S450L in *rpoB*, the RNAP β subunit (Table 2).

Drug resistance phenotypes and genotypes

Only three of the 33 East European Beijing isolates sequenced exhibited a pan-susceptible phenotype to the nine antibiotics tested (Table 3). Twenty-six were MDR, and 13 of these were classified as XDR. All isolates that were phenotypically resistant to isoniazid, rifampicin, streptomycin, and ethambutol carried mutations known to confer resistance (Zhang et al. 1992; Finken et al. 1993; Telenti et al. 1993, 1997; Banerjee et al. 1994).

The majority of pyrazinamide-resistant isolates harbored nonsynonymous (ns) SNPs within *pncA*, which is required for activation of the prodrug (Scorpio and Zhang 1996); however, not all *pncA* mutations were associated with resistance. Notably, the only *pncA* nsSNP that was not unique in the collection (I6L) did not confer resistance. Two isolates exhibited pyrazinamide resistance that could not be explained by mutation within *pncA* or its promoter.

Correlation of VNTR types with SNP clustering

Within the Beijing lineage, VNTR clustering showed a high degree of consistency with the SNP-derived phylogeny, exhibiting homoplasy at a single locus between VNTR types 24 and 29 (Fig. 3). Isolates with VNTR pattern 12 were the most diverse, exhibiting polymorphism at over 200 nucleotide positions. It is notable that for this VNTR pattern, the number of repeats at each locus is the most common number of repeats at that locus across the collection, suggesting that it represents the ancestral pattern of the East European isolates.

Within the East European group, the phylogeny revealed two clearly distinct clades of highly related isolates. The clades, designated A and B, each included isolates of several VNTR types. Within each of these clades, there were as few as 20 SNPs separating isolates with different VNTR patterns. Thus, isolates of different VNTR types can be more closely related than those that share a pattern.

Clade A included all sequenced members of the first- and fourth-most prevalent VNTR clusters in the population, types 4 and 17, which differ at a single VNTR locus, as well as two small clusters (less than 10 isolates). Assuming that all isolates belonging to at least these four VNTR clusters are members of clade

Figure 3. Correlation of VNTR types with SNP-derived phylogeny of the Beijing family. Branch labels show number of SNPs. VNTR loci are colored according to the prevalence of the number of repeats at that locus. Cluster designations are given for isolates from Samara; isolates with cluster number 0 were unique in the collection.

For five isolates, resistance to either or both of the ribosome-targeting antibiotics kanamycin and capreomycin could not be attributed to *rrs* or *eis* mutations (Maus et al. 2005b; Zaunbrecher et al. 2009). No isolates had mutations in *tlyA*, a gene disputably associated with capreomycin resistance (Maus et al. 2005a; Engström et al. 2011). To investigate possible novel drug resistance mutations, we searched for commonalities in SNPs within the terminal branches of these three strains, excluding those associated with nonaminoglycoside drug resistance and synonymous changes. The isolates did not share polymorphisms within any gene or promoter region, although the isolate resistant to both antibiotics carried a SNP within Rv3728, which has similarity to drug efflux proteins.

Eight of 14 isolates that were cross-resistant to the fluoroquinolones moxifloxacin and ofloxacin carried mutations in *gyrA*, which are known to confer resistance (Takiff et al. 1994). We investigated the co-occurrence of proximal SNPs in six isolates with wild-type *gyrA* sequences that independently acquired fluoroquinolone resistance, but we found no commonalities that might represent resistance mutations.

Evolution of drug resistance

Reconstruction of antibiotic-resistance genotypes and phenotypes onto the phylogeny of the Beijing isolates, assuming parsimonious acquisition with no reversion, enabled us to estimate the spread of primary resistance (Fig. 4).

In this model, the pan-susceptible ancestor of clade A, belonging to VNTR cluster 17, acquires rifampicin, isoniazid, and streptomycin resistance. The subsequent *eis* promoter and *embB* mutations, conferring kanamycin and ethambutol resistance, are coincident with a change in repeat locus 4052 to VNTR type 4 and with further clonal expansion. In contrast, acquisition of pyrazinamide and fluoroquinolone resistance arises only on terminal branches, occurring two and four times, respectively.

The expansion of clade B, from an isoniazid- and streptomycin-resistant ancestor is associated with development of rifampicin resistance conferring an MDR phenotype. Within this clade, both ethambutol and kanamycin resistance are independently acquired multiple times, including at least once by an inferred common ancestor. Again pyrazinamide and fluoroquinolone resistance was

Table 1. Homoplasies among Samaran Beijing isolates

Gene	Position	SNP	Substitution[a]	Associated resistance	Reference
gyrA	7,582	A>C	D94A	Fluoroquinolones	Takiff et al. (1994)
	7,582	A>G	D94G		
rpoB	761,100	T>C	L430P	Rifampicin	Telenti et al. (1993)
	761,160	C>T	S450L		
rpoC	764,846	T>C	I491T	—	—
rpsL	781,692	A>G	K43R	Streptomycin	Finken et al. (1993)
	781,827	A>G	K88R		
rrs	1,473,254	A>G	1401	Kanamycin/capreomycin	Maus et al. (2005b)
katG	2,155,176	C>G	S315T	Isoniazid	Zhang et al. (1992)
pncA	2,289,261	T>C	−11	Pyrazinamide	Scorpio and Zhang (1996)
eis	2,715,350	C>T	−10	Kanamycin	Zaunbrecher et al. (2009)
	2,715,352	G>A	−12		
	2,715,377	C>A	−37		
embA	4,243,224	C>T	−16	Ethambutol	Telenti et al. (1997)
	4,243,228	C>T	−12		
embB	4,247,436	A>G	M306V	Ethambutol	Telenti et al. (1997)
	4,247,438	G>A	M306I		
	4,248,010	A>G	Q497R		

[a]Amino acid substitutions are given for coding genes and nucleotide positions relative to gene start for noncoding genes and promoter regions.

only sporadically acquired, on terminal branches, occurring four and six times, respectively.

Discussion

The association of different *M. tuberculosis* lineages with different geographic populations has been described (Baker et al. 2004; Gagneux et al. 2006a). Here we show that geographical clustering is also observed within the Beijing lineage. Within the context of the global phylogeny of this lineage, represented by the isolates from London, 31 of the 32 Beijing isolates from Samara clustered together with the isolate representing the dominant clone in Estonia (Kruuner et al. 2001). This suggests that a single "East European" subtype of the Beijing lineage dominates this region. Comparison of our polymorphisms with those defining the Beijing K-type, dominant in Uzbekistan (Niemann et al. 2009), indicated that K-strains belong to the East European subtype but are not part of clades A or B.

Within the Samaran Beijing population, VNTR patterns appear relatively stable. The acquisition of drug resistance constitutes an evolutionary bottleneck that resembles that occurring in a transmission event. The Samaran isolates acquired resistance to up to four different drugs without concomitant VNTR change. Thus, in this high-incidence region, VNTR pattern identity between Beijing strains is insufficient to corroborate an epidemiological link and cannot be used as an indicator of recent transmission, nor can reactivation versus exogenous reinfection be distinguished.

Although VNTR patterns of EuroAmerican strains are generally more diverse (Brown et al. 2010), whole-genome sequencing of 36 isolates belonging to this lineage resolved an apparent outbreak from a single 24-loci VNTR type into two separate outbreaks associated with different clones (Gardy et al. 2011). Two further studies that attempted to document the evolution of MDR to XDR TB by whole-genome sequencing of isolates with identical fingerprints found that the isolates were not clonal (Ioerger et al. 2009; Niemann et al. 2009). These data indicate that VNTR typing is of limited use as an epidemiological tool in any region with a high-TB burden.

With one exception, the homoplasic SNPs identified were associated with drug resistance, indicating that antibiotic therapy exerts the primary evolutionary pressure on this *M. tuberculosis*

population. Accumulated evidence has revealed that in contrast to other bacterial pathogens, where antibiotic resistance is frequently acquired via horizontal transfer of resistance cassettes, *M. tuberculosis* resistance is mediated exclusively by chromosomal mutations affecting either the target itself or drug modifying enzymes (Zhang et al. 1992; Finken et al. 1993; Telenti et al. 1993, 1997; Banerjee et al. 1994; Takiff et al. 1994; Scorpio and Zhang 1996; Maus et al. 2005b; Zaunbrecher et al. 2009). This dogma is completely supported by our data.

Rifampicin resistance is overwhelmingly correlated with mutations within a short region of the RNAP β subunit, *rpoB*, which decrease drug binding affinity (Telenti et al. 1993). However, we were surprised to discover a homoplasy in the gene for the RNAP β′ subunit *rpoC*, which has never been associated with drug resistance in *M. tuberculosis*. Inspection of sequence data published by Ioerger et al. (2010a) revealed that of 14 South African Beijing isolates sequenced, one of eight isolates carrying *rpoB* S450L also carried the homoplasic *rpoC* mutation I491T. Again, the *rpoC* mutation was not found in rifampicin-sensitive isolates or those carrying alternate *rpoB*-resistance mutations.

Excluding one mutation (E1092C), which arose before the acquisition of rifampicin resistance, all nine nsSNPs in *rpoC* we identified were found exclusively in isolates carrying the *rpoB* S450L mutation, raising the possibility that they are compensatory mutations (Table 2). Intriguingly, mutations within a short region of *rpoA*, encoding the RNAP α subunit, correlated with the presence of the *rpoB* S450L mutation and absence of *rpoC* mutations, suggesting it could function as an alternative compensatory mechanism. In competitive growth assays, Gagneux et al. (2006b) found that the *rpoB* S450L mutation was associated with the lowest fitness cost in laboratory mutants and no fitness cost in four out of five clinical strains. It is possible that compensatory mutations in *rpoC* or *rpoA* are responsible for ameliorating the fitness defect in the clinical *rpoB* mutants.

Zaunbrecher et al. (2009) identified a mutation in the *eis* gene promoter as the source of spontaneously acquired kanamycin resistance in a laboratory *M. tuberculosis* strain. Sequencing the *eis* promoter region of clinical *M. tuberculosis* isolates has revealed a strong association with kanamycin resistance that is further supported by this study (Campbell et al. 2011; Huang et al. 2011). The promoter mutations result in increased expression of Eis, an aminoglycoside acetyl transferase that efficiently acetylates and inactivates kanamycin (Zaunbrecher et al. 2009; Chen et al. 2011). Wu et al. (2009) have shown that up-regulation of *eis*, in a recombinant Beijing strain, was associated with more rapid growth in human monocytes. Other studies have reported that Eis modulates secretion of pro-inflammatory cytokines, effectively suppressing a protective immune response and suggesting a critical role in infection outcome (Lella and Sharma 2007; Samuel et al. 2007; Shin et al. 2010). These effects of Eis expression raise the troubling possibility that kanamycin therapy may select for mutants with increased virulence, repudiating the general principle that drug resistance leads to decreased fitness. Mutations in the *eis* promoter may partly account for preservation of bacterial fitness

Casali et al.

Table 2. nsSNPs in RNA polymerase genes in Samaran Beijing isolates

| Position | Gene | Substitution | Reference | Nucleotide change[a] | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 08-14389 | 08-14365 | 08-14486 | 08-14361 | 08-00483 | 08-14306 | 08-14303 | 08-14337 | 08-14350 | 08-14521 | 08-14377 | 08-14556 | 08-14354 | 08-14329 | 08-14358 | 08-14493 | 08-14310 | 08-14567 | 08-14443 | 08-14569 | 08-14357 | 08-14477 | 08-14514 | 08-14546 | 08-14300 | 08-14391 | 08-14487 | 08-14370 | 08-14304 | 08-14543 | 08-14558 | 08-14559 | 08-14471 |
| 761,100 | rpoB | L430P | T | – | – | – | – | – | – | – | – | – | – | – | C | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 761,102 | | S431G | A | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 761,144 | | H445N | C | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 761,146 | | H445Q | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 761,160 | | S450L | C | T | – | T | T | T | T | T | T | T | – | T | – | T | T | – | T | – | – | – | – | – | – | T | T | T | T | T | T | T | T | T | T | T |
| 761,249 | | I480V | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 762,094 | | E761D | G | – | – | – | – | – | – | – | – | – | G | G | – | G | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – |
| 764,368 | rpoC | G332R | G | – | – | – | – | C | C | – | C | – | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – | C | C | C | C | C | C | C | C | C |
| 764,729 | | F452S | T | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 764,822 | | V483G | T | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 764,827 | | D485N | G | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 764,846 | | I491T | T | – | – | – | – | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 764,953 | | L527V | T | – | – | – | – | – | – | – | – | – | G | G | – | – | – | – | – | – | – | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – |
| 765,577 | | D735N | G | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 765,614 | | D747A | A | – | – | – | – | C | C | – | C | C | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 766,027 | | I885V | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | T | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – |
| 766,650 | | E1092D | A | – | – | – | – | – | – | – | – | – | – | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | C |
| 3,877,946 | rpoA | D190G | T | A | – | C | C | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 3,877,956 | | T187A | T | – | – | – | C | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |

[a]Dashes indicate that the nucleotide at this position matched the reference sequence.

Table 3. Correlation between drug resistance phenotypes and genotypes in Beijing isolates[a]

Isolate	VNTR cluster	RIF	rpoB[b]	INH	katG[b]	mabA	PZA	pncA	EMB	embA	embB	STR	rpsL	rrs[b]	CAP	KAN	rrs[b]	eis	MOX	OFL	gyrA[b]
08-14471	4	R	S450L	R	S315T	-	S	I6L	S	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14559	4	R	S450L	R	S315T	-	ND	I6L	R	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14558	4	R	S450L	R	S315T	-	S	I6L	R	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14543	4	R	S450L	R	S315T	-	ND	I6L	R	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14304	4	R	S450L	R	S315T	-	ND	I6L	R	-	D354A	R	-	517	S	R	-	-37	R	R	-
08-14370	4	R	S450L	R	S315T	-	S	I6L	R	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14487	6	R	S450L	R	S315T	-	R	A3E	R	-	D354A	R	-	517	S	R	-	-37	R	R	D94A
08-14391	4	R	S450L	R	S315T	-	R	W119R	R	-	D354A	R	-	517	S	R	-	-37	S	S	-
08-14300	4	R	S450L	R	S315T	-	ND	-	R	-12	D354A	R	-	517	S	R	-	-37	R	R	D94N
08-14546	42	R	S450L	R	S315T	-17	S	V44G	R	-	M306V	R	-	517	S	S	-	-12	R	R	D94G
08-14514	17	R	S450L	R	S315T	-	S	F58L	R	-	M306V	R	K43R	517	S	S	-	-14	S	S	-
08-14477	17	S	-	S	-	-	S	-	S	-	-	R	K43R	-	ND	ND	-	-	ND	ND	-
08-14357	17	S	-	S	-	-	ND	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
08-14569	0	S	-	R	S315T	-	ND	-	S	-	E504D	R	K88M	-	ND	ND	-	-	ND	ND	-
08-14443	12	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
08-14567	15	R	S450L	R	S315T	-	S	-	R	-12	M306V	R	K43R	-	S	S	-	-	S	S	-
08-14310	12	R	L430P/H445N	R	S315T	-15	ND	V139A	R	-	M306V	R	K88R	-	R	R	1401	-	R	R	D94G
08-14493	33	R	S450L	R	S315T	-	R	-11	R	-	M306V	R	K43R	-	S	R	-	-37	R	R	-
05-08961	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
08-14358	20	S	-	S	-	-	ND	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
08-14329	39	R	S450L	R	S315T	-	ND	S59P	R	-	H1002R	R	K43R	-	S	R	1401	-	S	S	-
08-14354	12	R	S450L	R	S315T	-	S	-	S	-	-	R	K88R	-	R	S	-	-	S	S	-
08-14556	15	R	L430P/S431G/H445Q	R	S315T	-	R	-	R	-	M306V	S	K43R	-	S	R	-	-37	R	R	A90V
08-14377	24	R	S450L	R	S315T	-	S	-	R	-	M306V	R	K43R	-	S	R	-	-10	R	R	-
08-14521	24	R	S450L	R	S315T	-	R	V128G	R	-	M306V	R	K43R	-	S	R	-	-	S	S	-
08-14350	29	R	S450L	R	S315T	-	ND	-11	R	-	M306V	R	K43R	-	S	R	-	-10	R	R	D94A
08-14337	29	R	S450L	R	S315T	-	ND	-	R	-	M306V	R	K43R	-	S	S	-	-10	S	S	-
08-14303	24	R	S450L	R	S315T	-	ND	-11	R	-	M306I	R	K43R	-	R	R	1401	-	R	R	-
08-14306	24	R	S450L	R	S315T	-	ND	D136Y	R	-	M306I/Q497R	R	K43R	-	R	R	-	-12	R	R	D94G
08-00483	24	R	S450L	R	S315T	-	S	-	R	-	M306V	R	K43R	-	S	R	-	-10	S	S	-
08-14361	25	R	S450L	R	S315T	-	S	T47A	S	-	N296H	R	K43R	-	S	R	-	-10	R	R	-
08-14486	29	R	S450L	R	S315T	-	S	A102V	R	-16	-	R	K43R	-	S	R	1401	-	R	R	-
08-14365	3	S	-	R	S315T	-	S	-	S	-16	-	R	K43R	-	S	S	-	-	S	S	-
05-03798	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
05-10261	-	S	-	S	-	-	S	-	S	-	-	R	K43R	-	ND	ND	-	-	ND	ND	-
05-06899	-	R	D435V	R	S315T	-	R	-11	R	-	P404S	R	K43R	-	S	S	-	-	S	S	-
05-03633	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
05-12611	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
05-09310	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
05-03802	-	S	-	S	-	-	S	-	S	-	-	S	-	-	ND	ND	-	-	ND	ND	-
06-01971	-	S	-	S	-	-	S	-	S	-	-	R	K43R	-	ND	ND	-	-	ND	ND	-
08-14389	0	R	S450L	R	S315T	-	R	T135P	R	-	Q497R	R	K43R	-	R	R	-	-	R	R	D94G

[a]Columns labeled RIF (rifampicin), INH (isoniazid), PZA (pyrazinamide), EMB (ethambutol), STR (streptomycin), CAP (capreomycin), KAN (kanamycin), MOX (moxifloxacin), and OFL (ofloxacin) show phenotypic susceptibility results: resistant (R), sensitive (S), or not determined (ND). Associated genotypes are shaded in the same color.
[b]Only known resistance-conferring mutations are shown for this gene.
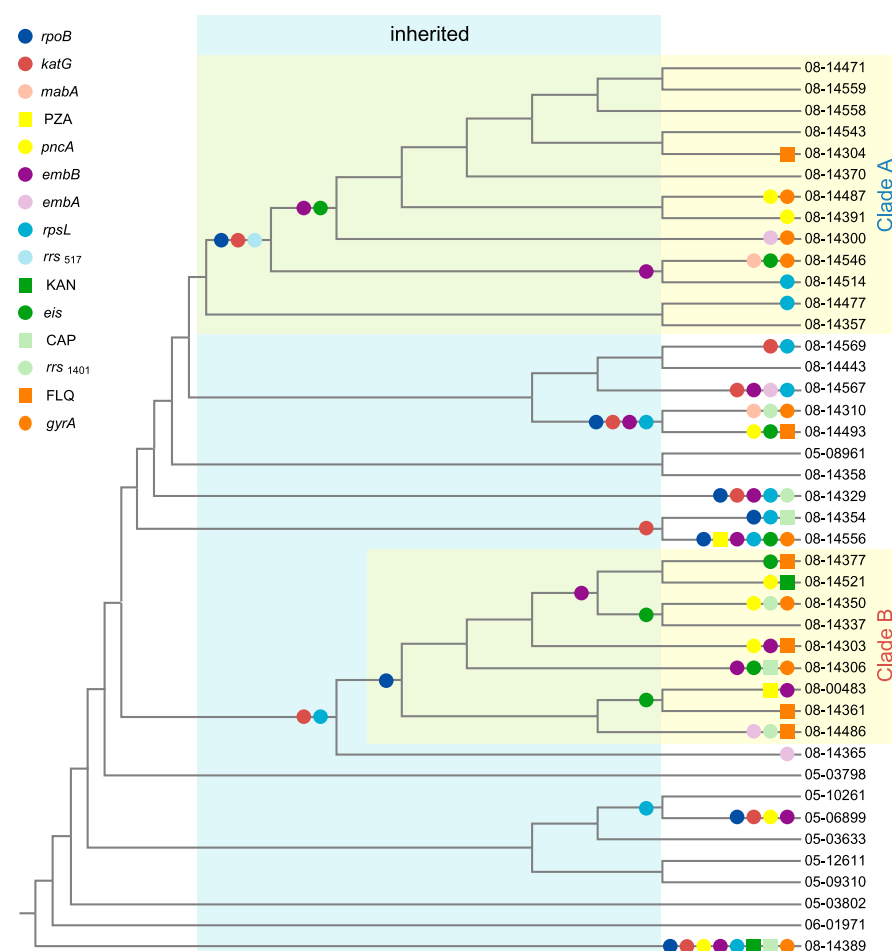
Figure 4. Model of acquisition of drug resistance in Beijing isolates. SNPs conferring drug resistance were mapped onto a cladogram of the Beijing family assuming parsimonious acquisition with no reversion (circles); SNPs are included if the majority of isolates carrying the mutation were phenotypically resistant (Table 3). Applying the same principle, resistance phenotypes with unknown genetic cause were added to the tree (squares; PZA indicates pyrazinamide; KAN, kanamycin; CAP, capreomycin; and FLQ, fluoroquinolones). Genotypes putatively acquired by an inferred common ancestor are boxed in blue.

and the extensive transmission of drug-resistant TB strains in Samara.

By mapping drug resistance genotypes and phenotypes onto the SNP-derived phylogenetic tree, we deduced that the prevalence of primary and acquired resistance varied widely between antibiotics. Cases of primary MDR TB were extremely common in the sequenced population. Expansion of clones with mutations in ethambutol-, streptomycin-, and kanamycin-resistance loci was also observed. In contrast, mutations in *pncA*, conferring resistance to the commonly used first-line drug pyrazinamide, were only found on terminal branches, suggesting acquired resistance. These data suggest that functional PncA is required to establish a new infection. Similarly, the appearance of fluoroquinolone resistance, typically conferring an XDR phenotype, was frequent but also sporadic, with no evidence of primary resistance. Although this may partly be explained by the more recent introduction of fluoroquinolone therapy, it is possible that the XDR mutants have impaired fitness for transmission.

Most studies that purport to describe transmission of XDR TB rely on molecular fingerprint data to infer clonality (for review, see Borrell and Gagneux 2009). However, XDR strains are

typically found in regions of high-TB prevalence where, we have shown, clonality can only be reliably determined through whole-genome sequencing. Two studies have described sequence-based investigations of XDR strains circulating in South Africa. Comparison of shared and unique polymorphisms in the genomes of 14 Beijing isolates suggested that although MDR strains were clonally transmitted, XDR was adaptively acquired (Ioerger et al. 2010a). Although the relatively small sample size of this study precluded wider generalization, the findings support our conclusions. In contrast, homogeneity between the genome sequences of eight XDR isolates belonging to the EuroAmerican KZN-type suggested clonal expansion of this strain (Ioerger et al. 2009). The isolates shared a *gyrA* mutation, indicating transmission of fluoroquinolone resistance; however, there was no evidence of primary pyrazinamide resistance. Further studies are required to determine whether there are genuine lineage-specific differences in the ability of drug-resistant strains to spread.

This study provides evidence that MDR clones have disseminated across Samara, supporting improved infection control as part of the wider strategy to address drug-resistant TB. Further large-scale sequencing projects will be necessary to track the emergence and spread of XDR TB to determine whether these strains begin to dominate the population or are too unfit for successful transmission. Understanding the variable fitness of different drug-resistant mutants will enable treatment regimens to be tailored both to meet the patients' needs and to achieve the greatest public health benefit.

Methods

MTBC isolates and preparation of genomic DNA

Isolates from the archives of the National *Mycobacterium* Reference Laboratory, London, United Kingdom, and the Samara TB Service, Samara Oblast, Russian Federation, were cultured on Lowenstein-Jensen slants for 4–6 wk at 37°C. Genomic DNA was isolated using a modification of the method described by Santos et al. (1992) and purified using a DNeasy Blood and Tissue kit (Qiagen).

Genomic library preparation and next-generation sequencing

Genomic libraries with 200-bp inserts were prepared for sequencing using standard protocols according to the method described by Harris et al. (2010). Libraries for multiplex sequencing, of 12 isolates per lane, included indexing tags. Singleplex or pooled indexed libraries were sequenced on an Illumina Genome Analyzer GAII at the Wellcome Trust Sanger Institute according

to the manufacturer's protocols for generating paired-end 54-bp reads.

Preparation of reference sequence

M. tuberculosis H37Rv was used as the reference sequence for mapping sequencing reads. We obtained the H37Rv strain that was originally sequenced from the Pasteur Institute (Cole et al. 1998), and resequenced the genome in this study.

By using iCORN (Iterative Correction of Reference Nucleotides) (Otto et al. 2010), we identified 56 SNPs and seventeen 1- or 2-bp indels relative to the published H37Rv sequence (AL123456 version 4). Sixty of these polymorphisms were detected in one or more previous H37Rv resequencing efforts (Zheng et al. 2008; Niemann et al. 2009; Ioerger et al. 2010b). The polymorphisms uncovered in these studies likely represent a combination of errors in the original reference sequence and genuine differences between H37Rv strains held in various laboratories. We used our corrected H37Rv sequence, which has a total length of 4,411,538 bp, as the reference genome in this study.

Annotation was downloaded from TubercuList (release R21; http://tuberculist.epfl.ch/) and transferred onto the corrected reference sequence. Locations of repeat regions were transferred from the EMBL sequence. The corrected annotated reference sequence is available online (Supplemental data 2).

Identification of SNPs

Illumina sequence reads for each isolate were mapped onto the H37Rv reference genome as paired-end reads with an insert size of 50–400 bp using SSAHA2 with default mapping parameters, except the minimum mapping score was increased to 30 (Ning et al. 2001). The mean mapping depth for seven *M. tuberculosis* isolates with singleplex data ranged from 337- to 582-fold (mean, 479), with 97.3%–98.8% (mean 97.8%) of the reference genome covered by reads. Data from multiplexed isolates, or combined data from two multiplex runs, were excluded from further analysis if the mean mapping depth was less than 15-fold. For 51 isolates with multiplex data from one or two runs, mean mapping depth ranged from 17- to 146-fold (mean, 56) and coverage of the reference genome from 93.8%–97.9% (mean, 96.3%).

Candidate SNPs were identified using ssaha_pileup (Ning et al. 2001). SNPs that were present in <75% of reads or with a quality score below 30 were filtered out. These parameters effectively eliminate false positives at the expense of false negatives (Harris et al. 2010). SNPs located in annotated repeat regions or within genes belonging to the Tuberculist categories "PE/PPE" or "insertion sequences and phages," which are notoriously difficult to map and/or sequence using current methodologies, were also excluded from further analysis. A total of 408,062 bp of the genome (9.25%), including 315 CDSs, was thus excluded.

A maximum likelihood phylogeny was reconstructed using a GTR model plus a gamma parameter with four categories to model for among site rate variation (Stamatakis et al. 2005). Calculation of 100 bootstrap replicates provided support for nodes on the tree. SNPs occurring on each branch of the tree were reconstructed using DELTRAN parsimony.

Identification of LSPs

Sites of large deletions (>500 bp) were conservatively inferred by comparison of mapping coverage plots. Where regions had zero coverage in all strains, this was attributed to sequencing or mapping artifacts; regions where no reads mapped in only a subset of strains were considered to represent genuine deletions. For East European Beijing isolates, unmapped reads were assembled de novo using Velvet (Zerbino and Birney 2008). Contigs longer than 500 bp were mapped against completed *M. tuberculosis* genome sequences (NCBI genomic BLAST; http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi).

Molecular fingerprinting

Isolates were characterized by spoligotyping according to standard methods (Kamerbeek et al. 1997). VNTR typing was conducted using a panel of 15 loci (Kwara et al. 2003) or, for isolates with a Beijing spoligotype, with a panel of eight MIRU-VNTR loci (MIRU10, 23, 26, 31, 40; VNTR 3232, 1982, 4052) (Nikolayevskyy et al. 2006). VNTR typing was performed in Samara by monoplex PCR and agarose gel electrophoresis and in London using a high-throughput method based on multiplex PCR with fluorescent-labeled primers followed by fragment separation on a CEQ8000 Genome Instrument (Beckman Coulter) (Brown et al. 2009).

Drug susceptibility testing

Drug susceptibility to the first-line drugs rifampicin, isoniazid, streptomycin, and ethambutol was performed using the automated Mycobacterial Growth Indicator Tube (MGIT) 960 system (Becton Dickinson) (Kruuner et al. 2006). For MDR isolates, susceptibility to the second-line drugs kanamycin, capreomycin ofloxacin, and moxifloxacin was also determined (Kruuner et al. 2006). Pyrazinamide susceptibility was determined in the MGIT system in Samara or in semisolid media in London (Collins et al. 1997).

Data access

Raw sequence data have been submitted to the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under accession no. ERP000192.

Acknowledgments

References

Baker L, Brown T, Maiden MC, Drobniewski F. 2004. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg Infect Dis* **10:** 1568–1577.

Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um K, Wilson T, Collins D, de Lisle G, Jacobs W. 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* **263:** 227–230.

Borrell S, Gagneux S. 2009. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* **13:** 1456–1466.

Brown TJ, Nikolayevskyy VN, Drobniewski FA. 2009. Typing *Mycobacterium tuberculosis* using variable number tandem repeat analysis. *Methods Mol Biol* **465:** 371–394.

Brown T, Nikolayevskyy V, Velji P, Drobniewski F. 2010. Associations between *Mycobacterium tuberculosis* strains and phenotypes. *Emerg Infect Dis* **16:** 272–280.

Campbell PJ, Morlock GP, Sikes RD, Dalton TL, Metchock B, Starks AM, Hooks DP, Cowan LS, Plikaytis BB, Posey JE. 2011. Molecular detection of mutations associated with first- and second-line drug resistance compared with conventional drug susceptibility testing of *Mycobacterium tuberculosis. Antimicrob Agents Chemother* **55:** 2032–2041.

Chen W, Biswas T, Porter VR, Tsodikov OV, Garneau-Tsodikova S. 2011. Unusual regioversatility of acetyltransferase Eis, a cause of drug resistance in XDR-TB. *Proc Natl Acad Sci* **108:** 9804–9808.

Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393:** 537–544.

Collins CH, Grange JM, Yates MD. 1997. Drug susceptibility testing. In *Tuberculosis bacteriology: Organization and practice*, pp. 98–109. Butterworth-Heinemann, Oxford, UK.

Drobniewski F, Balabanova Y, Nikolayevsky V, Ruddy M, Kuznetzov S, Zakharova S, Melentyev A, Fedorin I. 2005. Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia. *JAMA* **293:** 2726–2731.

Engström A, Perskvist N, Werngren J, Hoffner SE, Juréen P. 2011. Comparison of clinical isolates and in vitro selected mutants reveals that *tlyA* is not a sensitive genetic marker for capreomycin resistance in *Mycobacterium tuberculosis. J Antimicrob Chemother* **66:** 1247–1254.

Finken M, Kirschner P, Meier A, Wrede A, Böttger EC. 1993. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Mol Microbiol* **9:** 1239–1246.

Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, et al. 2006a. Variable host–pathogen compatibility in *Mycobacterium tuberculosis. Proc Natl Acad Sci* **103:** 2869–2873.

Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. 2006b. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis. Science* **312:** 1944–1946.

Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, Zeller K, Andrews J, Friedland G. 2006. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **368:** 1575–1580.

Gardy JL, Johnston JC, Sui SJH, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R, et al. 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* **364:** 730–739.

Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, et al. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327:** 469–474.

Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM. 2004. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci* **101:** 4871–4876.

Huang W-L, Chi T-L, Wu M-H, Jou R. 2011. Performance assessment of the GenoType MTBDR*sl* test and DNA sequencing for detection of second-line and ethambutol drug resistance among patients infected with multidrug-resistant *Mycobacterium tuberculosis J Clin Microbiol*. **49:** 2502–2508.

Ioerger TR, Koo S, No E-G, Chen X, Larsen MH, Jacobs WR, Pillay M, Sturm AW, Sacchettini JC. 2009. Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PLoS ONE* **4:** e7778. doi: 10.1371/journal.pone.0007778.

Ioerger TR, Feng Y, Chen X, Dobos KM, Victor TC, Streicher EM, Warren RM, Gey van Pittius NC, Van Helden PD, Sacchettini JC. 2010a. The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. *BMC Genomics* **11:** 670. doi: 10.1186/1471-2164-11-670.

Ioerger TR, Feng Y, Ganesula K, Chen X, Dobos KM, Fortune S, Jacobs WR, Mizrahi V, Parish T, Rubin E, et al. 2010b. Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories. *J Bacteriol* **192:** 3645–3653.

Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, et al. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* **35:** 907–914.

Kliiman K, Altraja A. 2009. Predictors of poor treatment outcome in multi- and extensively drug-resistant pulmonary TB. *Eur Respir J* **33:** 1085–1094.

Kruuner A, Hoffner SE, Sillastu H, Danilovits M, Levina K, Svenson SB, Ghebremichael S, Koivula T, Kallenius G. 2001. Spread of drug-resistant pulmonary tuberculosis in Estonia. *J Clin Microbiol* **39:** 3339–3345.

Kruuner A, Yates MD, Drobniewski FA. 2006. Evaluation of MGIT 960-based antimicrobial testing and determination of critical concentrations of first- and second-line antimicrobial drugs with drug-resistant

clinical strains of *Mycobacterium tuberculosis. J Clin Microbiol* **44:** 811–818.

Kwara A, Schiro R, Cowan LS, Hyslop NE, Wiser MF, Roahen HS, Kissinger P, Diem L, Crawford JT. 2003. Evaluation of the epidemiologic utility of secondary typing methods for differentiation of *Mycobacterium tuberculosis* isolates. *J Clin Microbiol* **41:** 2683–2685.

Leimane V, Dravniece G, Riekstina V, Sture I, Kammerer S, Chen MP, Skenders G, Holtz TH. 2010. Treatment outcome of multidrug/extensively drug-resistant tuberculosis in Latvia, 2000–2004. *Eur Respir J* **36:** 584–593.

Lella RK, Sharma C. 2007. Eis (enhanced intracellular survival) protein of *Mycobacterium tuberculosis* disturbs the cross regulation of T-cells. *J Biol Chem* **282:** 18671–18675.

Maus CE, Plikaytis BB, Shinnick TM. 2005a. Mutation of *tlyA* confers capreomycin resistance in *Mycobacterium tuberculosis. Antimicrob Agents Chemother* **49:** 571–577.

Maus CE, Plikaytis BB, Shinnick TM. 2005b. Molecular analysis of cross-resistance to capreomycin, kanamycin, amikacin, and viomycin in *Mycobacterium tuberculosis. Antimicrob Agents Chemother* **49:** 3192–3197.

Migliori GB, Lange C, Centis R, Sotgiu G, Mütterlein R, Hoffmann H, Kliiman K, De Iaco G, Lauria FN, Richardson MD, et al. 2008. Resistance to second-line injectables and treatment outcomes in multidrug-resistant and extensively drug-resistant tuberculosis cases. *Eur Respir J* **31:** 1155–1159.

Nicol MP, Wilkinson RJ. 2008. The clinical consequences of strain diversity in *Mycobacterium tuberculosis. Trans R Soc Trop Med Hyg* **102:** 955–965.

Niemann S, Köser CU, Gagneux S, Plinke C, Homolka S, Bignell H, Carter RJ, Cheetham RK, Cox A, Gormley NA, et al. 2009. Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints. *PLoS ONE* **4:** e7407. doi: 10.1371/journal.pone.0007407.

Nikolayevskyy V, Gopaul K, Balabanova Y, Brown T, Fedorin I, Drobniewski F. 2006. Differentiation of tuberculosis strains in a population with mainly Beijing-family strains. *Emerg Infect Dis* **12:** 1406–1413.

Ning Z, Cox AJ, Mullikin JC. 2001. SSAHA: A fast search method for large DNA databases. *Genome Res* **11:** 1725–1729.

Otto TD, Sanders M, Berriman M, Newbold C. 2010. Iterative correction of reference nucleotides (iCORN) using second generation sequencing technology. *Bioinformatics* **26:** 1704–1707.

Samara Oblast Tuberculosis Dispensary. 2009. *Annual Report on TB Control in Samara Oblast*. Samara Oblast, Russia.

Samuel LP, Song C-H, Wei J, Roberts EA, Dahl JL, Barry CE, Jo E-K, Friedman RL. 2007. Expression, production and release of the Eis protein by *Mycobacterium tuberculosis* during infection of macrophages and its effect on cytokine secretion. *Microbiology* **153:** 529–540.

Santos AR, de Miranda AB, Lima LM, Suffys PN, Degrave WM. 1992. Method for high yield preparation in large and small scale of nucleic acids from mycobacteria. *J Microbiol Methods* **15:** 83–94.

Schürch AC, Kremer K, Warren RM, Hung NV, Zhao Y, Wan K, Boeree MJ, Siezen RJ, Smith NH, van Soolingen D. 2011. Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the *Mycobacterium tuberculosis* Beijing genotype. *Infect Genet Evol* **11:** 587–597.

Scorpio A, Zhang Y. 1996. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med* **2:** 662–667.

Shah NS, Pratt R, Armstrong L, Robison V, Castro KG, Cegielski JP. 2008. Extensively drug-resistant tuberculosis in the United States, 1993–2007. *JAMA* **300:** 2153–2160.

Shin D-M, Jeon B-Y, Lee H-M, Jin HS, Yuk J-M, Song C-H, Lee S-H, Lee Z-W, Cho S-N, Kim J-M, et al. 2010. *Mycobacterium tuberculosis* Eis regulates autophagy, inflammation, and cell death through redox-dependent signaling. *PLoS Pathog* **6:** e1001230. doi: 10.1371/journal.ppat.1001230.

Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21:** 456–463.

Takiff HE, Salazar L, Guerrero C, Philipp W, Huang WM, Kreiswirth B, Cole ST, Jacobs WR, Telenti A. 1994. Cloning and nucleotide sequence of *Mycobacterium tuberculosis gyrA* and *gyrB* genes and detection of quinolone resistance mutations. *Antimicrob Agents Chemother* **38:** 773–780.

Telenti A, Imboden P, Marchesi F, Matter L, Schopfer K, Bodmer T, Lowrie D, Colston M, Cole S. 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis. Lancet* **341:** 647–651.

Telenti A, Philipp WJ, Sreevatsan S, Bernasconi C, Stockbauer KE, Wieles B, Musser JM, Jacobs WR. 1997. The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat Med* **3:** 567–570.

Vitol I, Driscoll J, Kreiswirth B, Kurepina N, Bennett KP. 2006. Identifying *Mycobacterium tuberculosis* complex strain families using spoligotypes. *Infect Genet Evol* **6:** 491–504.

Wright A, Zignol M, Van Deun A, Falzon D, Gerdes SR, Feldman K, Hoffner S, Drobniewski F, Barrera L, van Soolingen D, et al. 2009. Epidemiology of antituberculosis drug resistance 2002–07: An updated analysis of the global project on anti-tuberculosis drug resistance surveillance. *Lancet* **373:** 1861–1873.

Wu S, Barnes PF, Samten B, Pang X, Rodrigue S, Ghanny S, Soteropoulos P, Gaudreau L, Howard ST. 2009. Activation of the *eis* gene in a W-Beijing strain of *Mycobacterium tuberculosis* correlates with increased SigA levels and enhanced intracellular growth. *Microbiology* **155:** 1272–1281.

Zaunbrecher MA, Sikes RD, Metchock B, Shinnick TM, Posey JE. 2009. Overexpression of the chromosomally encoded aminoglycoside acetyltransferase *eis* confers kanamycin resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci* **106:** 20004–20009.

Zerbino DR, Birney E. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* **18:** 821–829.

Zhang Y, Heym B, Allen B, Young D, Cole S. 1992. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **358:** 591–593.

Zheng H, Lu L, Wang B, Pu S, Zhang X, Zhu G, Shi W, Zhang L, Wang H, Wang S, et al. 2008. Genetic basis of virulence attenuation revealed by comparative genomic analysis of *Mycobacterium tuberculosis* strain H37Ra versus H37Rv. *PLoS ONE* **3:** e2375. doi: 10.1371/journal. pone.0002375.

Microevolution of extensively drug-resistant tuberculosis in Russia

Nicola Casali, Vladyslav Nikolayevskyy, Yanina Balabanova, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2012/02/01/gr.128678.111.DC1
References	This article cites 53 articles, 23 of which can be accessed free at: http://genome.cshlp.org/content/22/4/735.full.html#ref-list-1
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**