A rapid assay for detection of the epidemiologically important Central Asian/Russian strain of *Mycobacterium tuberculosis* Beijing genotype

I. Mokrousov, E. Chernyaeva, A. Vyazovaya, Y. Skiba, N. Solovieva, V. Valcheva, K. Levina, N. Malakhova, W. W. Jiao, L. L. Gomes, P. Suffys, M. Kütt, N. Aitkhozhina, A. D. Shen, O. Narvskaya, V. Zhuravlev

Supplementary material

Table S1. M. tuberculosis strain collections used for experimental validation of the PCR-RFLP assay.

Country (ref.) **Genotype**	Russia (Vyazovaya et al., 2015)	Estonia (Vyazovaya et al., 2017)	Bulgaria (Valcheva et al., 2008)	Kazakhstan (Skiba et al., 2015)	China (Jiao et al., 2008)	Vietnam (Mokrousov et al., 2006)	Brazil (Gomes et al., 2015)
Beijing, all	**77**			**49**	**41**		**10**
Beijing 94-32-cluster (Central Asian/Russian strain)	38			45			
Beijing B0/W148 (100-32)-cluster (successful Russian strain)	30			3			
Beijing other	9			1	41		10
Non-Beijing, all	**21**	**42**	**46**	**29**		**27**	
LAM	4	12	3	13		1	
Ural	7	10	1	8			
Haarlem	1	9	9	2		1	
NEW-1				3			
S	1		15				
T (SIT53)	8	8	14			2	
X		2	1				
TUR			3	1			
Cameroon				1			
CAS				1			
EAI						23	
M. africanum		1					

Beijing clusters were defined by VNTR typing (24-MIRU-VNTR loci scheme of Supply et al., 2006). Non-Beijing families were defined by spoligotyping followed by expert-based correction (LAM, NEW-1, Ural, TUR, Cameroon families) and testing specific SNPs (LAM and S families).

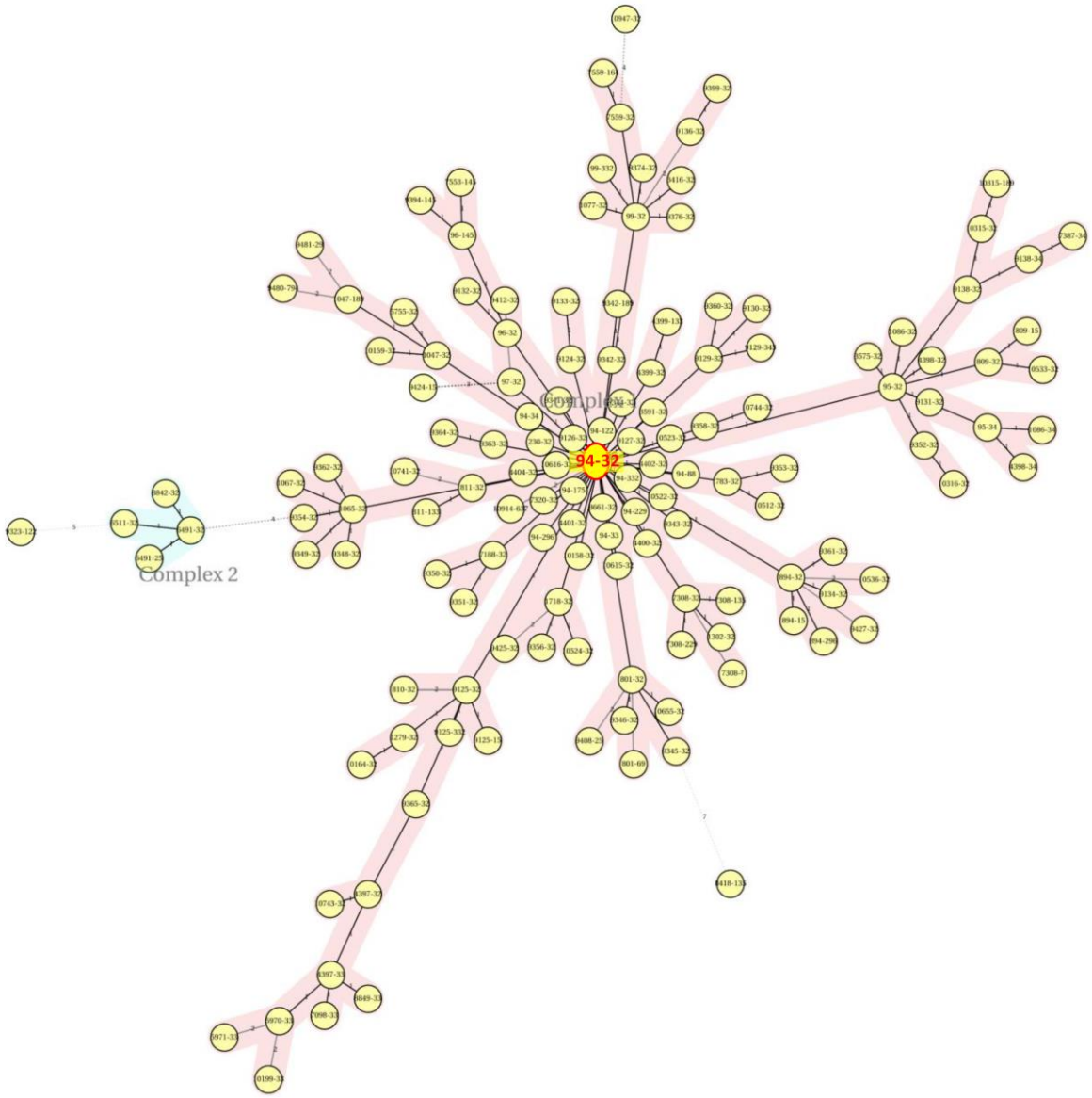

Figure S1. 24-MIRU-VNTR minimum spanning tree of the global dataset of the CC1 clonal complex isolates of *M. tuberculosis* taken from Merker et al. (2015), built with MIRU-VNTRplus.org.

Circle size is not to scale; the core type 94-32 includes 56% of all isolates.
The CC1 clonal complex defined by 24-MIRU-VNTR (Merker et al., 2015) correlates with WGS-defined East Europe 1 group (Luo et al., 2015) and Central Asian/Russian strain.

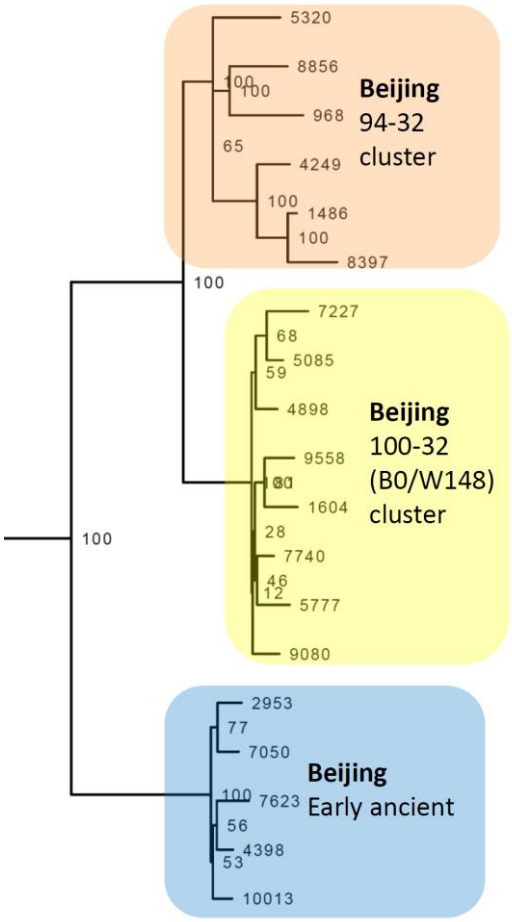

Figure S2. ML tree of 19 *M. tuberculosis* Beijing genotype isolates, based on NGS/WGS data.

The 19 isolates represented both modern and ancient sublineages and included six isolates of 94-32-cluster, eight isolates of 100-32-cluster and five isolates of the early ancient sublineage (deleted RD181, *mutT4* codon 48 wild type).

The quality of *M. tuberculosis* sequence reads was evaluated using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequence reads were processed using Trimmomatic (Bolger et al., 2014) and used for further nucleotide variation analysis. *M. tuberculosis* H37Rv reference genome (GenBank accession number NC_000962.3) was used for SNP calling. Sequencing reads were aligned on reference genome by using the bowtie2 program (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) and for further SNP calling and variant cell format file processing by using a combination of SAMtools, bcftools (http://samtools.sourceforge.net) and VCFtools (http://vcftools.sourceforge.net). Obtained VCF files were used for comprehensive genetic analysis. Single nucleotide polymorphisms (SNPs) with quality scores ≥20 were used for phylogenetic analysis. SNPs located in repetitive genome regions and in PE and PPE genes were removed from the analysis due to possible misalignments. Concatenated SNPs were used for phylogenetic analysis with RaxML package using maximum likelihood approach with GTRCAT

model and 100 bootstrap replications (Stamatakis, 2006) and using the *Mycobacterium canettii* genome as an outgroup.

The identified SNP in *sigE* gene codon 98 CTG>CTA (position 1364706 G>A in genome of reference strain H37Rv NC_000962.3) specific for the Beijing 94-32-cluster, was searched in the GMTV Database (http://mtb.dobzhanskycenter.org/) that included information on 2501 genomes at the time of comparison. We identified 537 strains with this mutation in the GMTV database and among those, 525 were designated as the Beijing lineage. The remaining 12 strains were labeled as 'New-1, Uganda, X-type' (n=6), 'not defined' (n=5), and Ural (n=1). An additional analysis of the vcf files with PhyTB online tool (http://pathogenseq.lshtm.ac.uk/phytblive/index.php) revealed that all these 12 samples presented a mixture of Lineage 2 and Lineage 4 genomes, and apparently, Lineage 2 (Beijing genotype) DNA contained a mutant allele of the studied *sigE* SNP. It was not possible to further correlate this mutation with MIRU-VNTR type of isolate since this feature is not included in the GMTV database and otherwise cannot be deduced from short sequencing reads.

Supplementary References

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma Oxf Engl 30:2114–2120.

Gomes LL, Vasconcellos SE, Gomes HM, Elias AR, da Silva Rocha A, Ribeiro SC, Panunto AC, Ferrazoli L, da Silva Telles MA, Ivens de AM, Kritski AL, Mokrousov I, Manicheva OA, Lasunskaia E, Suffys PN. Genetic diversity of the Mycobacterium tuberculosis Beijing family in Brazil and Mozambique and relation with infectivity and induction of necrosis in THP-1 cells. Tuberculosis (Edinb). 2015;95 Suppl 1:S190-6.

Jiao WW, Mokrousov I, Sun GZ, Guo YJ, Vyazovaya A, Narvskaya O, Shen AD. Evaluation of new variable-number tandem-repeat systems for typing Mycobacterium tuberculosis with Beijing genotype isolates from Beijing, China. J Clin Microbiol. 2008; 46:1045-9.

Mokrousov I, Jiao WW, Valcheva V, Vyazovaya A, Otten T, Ly HM, Lan NN, Limeschenko E, Markova N, Vyshnevskiy B, Shen AD, Narvskaya O. Rapid detection of the Mycobacterium tuberculosis Beijing genotype and its ancient and modern sublineages by IS6110-based inverse PCR. J Clin Microbiol. 2006;44:2851-6.

Skiba Y, Mokrousov I, Ismagulova G, Maltseva E, Yurkevich N, Bismilda V, Chingissova L, Abildaev T, Aitkhozhina N. Molecular snapshot of Mycobacterium tuberculosis population in Kazakhstan: a country-wide study. Tuberculosis (Edinb). 2015; 95:538-46.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. J Clin Microbiol 44:4498-4510.

Valcheva V, Mokrousov I, Narvskaya O, Rastogi N, Markova N. Molecular snapshot of drug-resistant and drug-susceptible Mycobacterium tuberculosis strains circulating in Bulgaria. Infect Genet Evol . 2008;8:657-63.

Vyazovaya A., Levina K., Zhuravlev V., Viiklepp P., Kütt M., Mokrousov I. Emerging resistant clones of *Mycobacterium tuberculosis* in spatiotemporal context. J Antimicrob Chemother DOI: 10.1093/jac/dkx372 in press.

Vyazovaya, A. Mokrousov I, Solovieva N, Mushkin A, Manicheva O, Vishnevsky B, Zhuravlev V, Narvskaya O. Tuberculous spondylitis in Russia and prominent role of multidrug-resistant clone Mycobacterium tuberculosis Beijing B0/W148. Antimicrob. Agents Chemother. 2015;59:2349-57.

Luo T, Comas I, Luo D, Lu B, Wu J, Wei L, Yang C, Liu Q, Gan M, Sun G, Shen X, Liu F, Gagneux S, Mei J, Lan R, Wan K, Gao Q. 2015. Southern East Asian origin and coexpansion of Mycobacterium tuberculosis Beijing family with Han Chinese. Proc Natl Acad Sci U S A 112:8136-8141.

Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, Blum MG, Rüsch-Gerdes S, Mokrousov I, Aleksic E, Allix-Béguec C, Antierens A, Augustynowicz-Kopeć E, Ballif M, Barletta F, Beck HP, Barry CE 3rd, Bonnet M, Borroni E, Campos-Herrero I, Cirillo D, Cox H, Crowe S, Crudu V, Diel R, Drobniewski F, Fauville-Dufaux M, Gagneux S, Ghebremichael S, Hanekom M, Hoffner S, Jiao WW, Kalon S, Kohl TA, Kontsevaya I, Lillebæk T, Maeda S, Nikolayevskyy V, Rasmussen M, Rastogi N, Samper S, Sanchez-Padilla E, Savic B, Shamputa IC, Shen A, Sng LH, Stakenas P, Toit K, Varaine F, Vukovic D, Wahl C, Warren R, Supply P, Niemann S, Wirth T. 2015. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. Nat Genet 47:242-249.