NAZARBAYEV UNIVERSITY
NATIONAL LABORATORY ASTANA

WHOLE GENOME SEQUENCING OF *M. TUBERCULOSIS* STRAINS IN KAZAKHSTAN REVEAL GENOMIC VARIANTS IN GENES CODING PE/PPE PROTEIN FAMILY SPECIFIC FOR MDR/XDR ISOLATES

U. Kozhamkulov[1], U. Kairov[2], A. Molkenov[2], M.Steins[3], A. Akhmetova[1], A.Akilzhanova[1]

[1]Laboratory of Genomic and Personalized Medicine, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan

[2]Laboratory of Bioinformatics and Systems Biology, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan

[3] Roche Diagnostics Deutschland GmbH, Mannheim, Germany

Introduction

➢ Worldwide in 2016, around **10.4 million** people were diagnosed with TB among which **1.7 million** died [1].

➢Despite the progress in decreasing the global incidence of drug-susceptible TB, multidrug-resistant (**MDR**) and extensively drug-resistant (**XDR**) tuberculosis in the past decade led to decreased efficiency of chemotherapy. However, the rate of drug-resistant TB increases annually, especially multidrug-resistant TB (MDR TB) [1].

➢ In Kazakhstan, the incidence of TB in 2016 was **52.7** cases per 100,000 and the mortality rate was **3.4** per 100,000 [2].

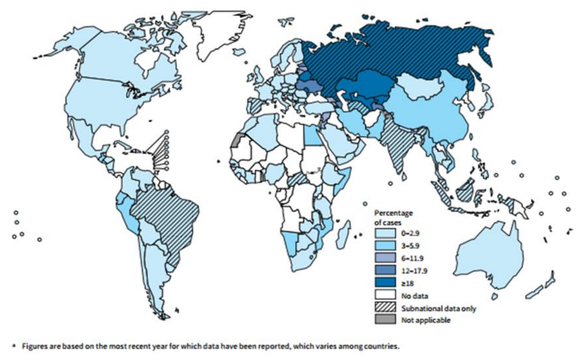
➢ The incidences of drug resistance, especially **multidrug-resistant (MDR) TB**, are increasing last 5 years in Kazakhstan [2].

➢ **27 countries** relate to **high MDR-TB** burden countries according WHO data, 14 of them in the European region, including **Kazakhstan** [1].

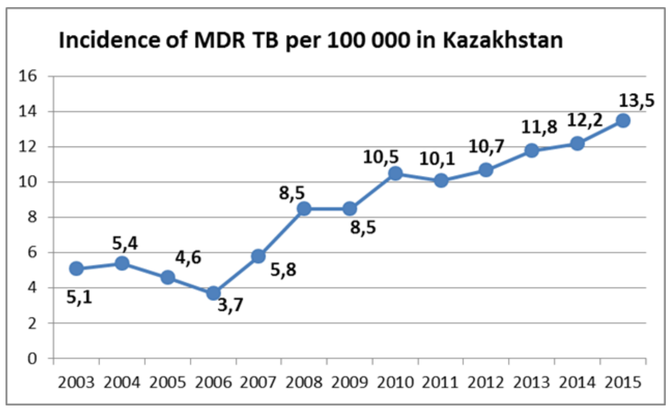
➢ At the moment in many countries around the world noted the spread of **Beijing family** strains of *M.tuberculosis* which associated with a high risk of drug resistance.

➢ Beijing/W of *Mycobacterium tuberculosis* strains are **the most virulent and prevalent** among young individuals.

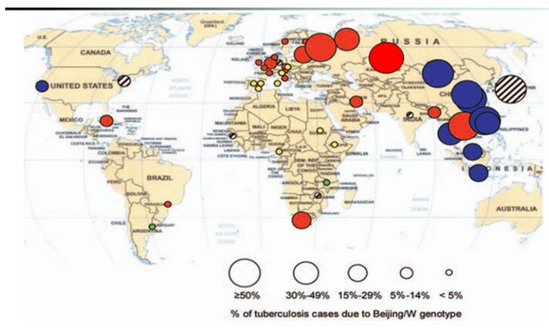
➢ This genotype is **widely distributed** in many parts of the world, including areas where it was not found previously [3].



Percentage of new MDR TB cases around the world. Kazakhstan is among countries with the highest MDR percentage [1]



Incidence of MDR TB per 100 000 in Kazakhstan [2] (National TB Center data)



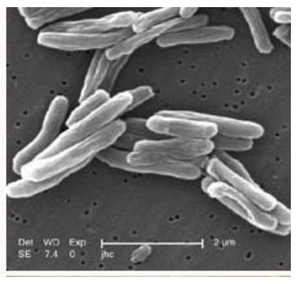
Distribution of Beijing genotype of *M.tuberculosis*

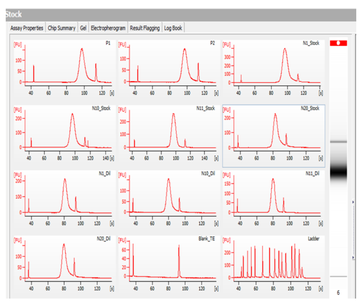
Emerging Infectious Diseases www.cdc.gov/eid • Vol. 12, No. 5, May 2008

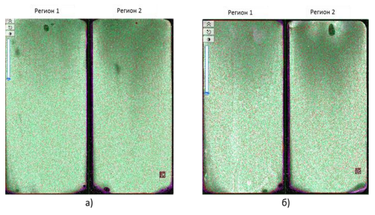
Aim

➢ The aim of this study is whole genome sequencing of *M.tuberculosis* clinical isolates with different drug susceptibility, the study of genetic markers of drug resistant TB and comparative analysis between the *M.tuberculosis* strains

Materials and Methods







➢ Clinical isolates of *M. tuberculosis* were collected from TB patients at National TB Center of the Republic of Kazakhstan, Almaty

➢ Drug susceptibility testing of *M.tuberculosis* was determined by absolute concentration method on solid Lowenstein-Jensen medium and by using the BACTEC-MGIT 960 Mycobacteria Growth Indicator Tube (BD Diagnostic Systems, USA) system

➢ Spoligotyping was performed on DNA by using the standard method [4] using a reverse dot-blot spoligotyping commercially available kit

➢ The whole genome sequencing of 20 *M.tuberculosis* clinical isolates from TB patients with different drug resistance (8 MDR, 3XDR, 1 mono-resistant, 1 poly-resistant, 7 susceptible) were performed by NGS platform Roche 454 GS FLX+ according to standard protocols

➢ Complete genomes assembling was performed by using NEWBLER de novo assembler (454 Life Sciences, Branford, CT)

➢ Alignment and mapping of sequences was performed on reference strain of *M.tuberculosis* H37Rv (NC_000962.3, GCF_000195955.2) by using the GS Reference Mapping (454 Life Sciences, Branford, CT)

➢ Comparative analysis of genomic variants among isolates was performed using Venn diagram

Results

➢ Most of the studied isolates of *M. tuberculosis* (18) are W-Beijing family strains (East Asian), only two isolates belongs to families T (Euro-American) and MANU-1 (Indo-Oceanic)

➢ International «Tuberculosis Drug Resistance Mutation Database» was analyzed to study the genetic loci involved in the resistance to basic anti-TB drugs. As a result of the analysis the most common mutations in 37 genes associated with drug resistance to 9 basic anti-TB drugs were selected

➢ Genetic mutations in genes associated with drug resistance of *M.tuberculosis* were analyzed. All detected genomic variants with single nucleotide polymorphisms, insertions, deletions for each clinical isolates of *M.tuberculosis* indicating the description of the gene and protein, the positions on reference genome H37Rv, deep coverage of gene were prepared

➢ Three main groups for comparative bioinformatics analysis were chosen – susceptible, MDR and XDR. 1018 genomic loci were identified as a common for all three study groups. The major parts of these genomic variants are found in “core” genes that necessary for mycobacteria life-sustaining activity

➢ 45 and 33 unique genetic loci were found between MTB-07-002 and MTB-06-003, also between MTB-07-002 and MTB-07-006, respectively

➢ Several genomic variants have been detected in four genes PE_PGRS24, PPE24, PPE5, PE_PGRS56 which are typical for MDR and XDR isolates and belong to genes of protein family PE/PPE specific for species of Mycobacteria only

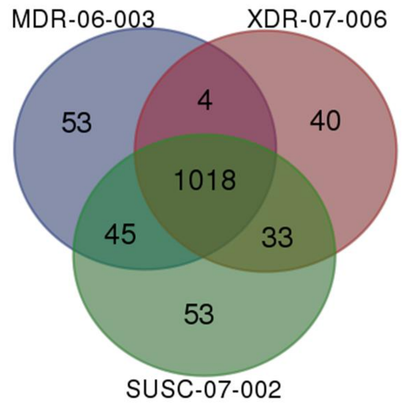
➢ The mutations specific to MDR and XDR groups of *M.tuberculosis* isolates can be one of the additional virulence factors that may provide an advantage in host-pathogen interaction

The main loci of *M.tuberculosis* drug resistance according to «Tuberculosis Drug Resistance Mutation Database»»

Anti-TB drugs	Gene locus	Gene
AMINOGLYCOSIDES (KANAMYCIN / CAPREOMYCIN / AMIKACIN / VIOMYCIN)	MTB000019, Rv1694	rrs, tlyA
ETHAMBUTOL	Rv0340, Rv0341, Rv0342, Rv0343, Rv1267c, Rv3124, Rv3125c, Rv3126, Rv3264c, Rv3266c, Rv3793, Rv3794, Rv3795	iniB, iniA, iniC, embR
ETHIONAMIDE	Rv1483, Rv1484, Rv3854c	mabA, inhA, ethA
FLUOROQUINOLONES	Rv0005, Rv0006	gyrB, gyrA
ISONIAZID	Rv0129c, Rv0340, Rv0341, Rv0342, Rv0343, Rv1483, Rv1484, Rv1592c, Rv1772, Rv1854c, Rv1908c, Rv1909c, Rv2242, Rv2243, Rv2245, Rv2247, Rv2427a, Rv2428, Rv2846c, Rv3139, Rv3566c, Rv3795	fbpC, iniB, iniA, iniC, mabA, inhA, ndh, katG, furA, srmR, fabD, kasA, accD6, oxyR, aphC, efpA, fadE24, nhoA, embB
PARA-AMINOSALISYLIC ACID	Rv2764c	thyA
PYRAZINAMIDE	Rv2043c	pncA
RIFAMPICIN	Rv0667, Rv3795	rpoB, embB
STREPTOMYCIN	MTB000019, Rv0682, Rv3919c	rrs, rpsL, gidB

Example of detected genomic variants for clinical isolate MTB-07-007

MTB-07-007					
Locus	Gene	Protein / product	The starting position in the reference genome	Nucleotide	Genomic variant
Rv0006	gyrA	DNA gyrase subunit A	7582	A	C
Rv0006	gyrA	DNA gyrase subunit A	7585	G	C
Rv0006	gyrA	DNA gyrase subunit A	9304	G	A
Rv0006	gyrA	DNA gyrase subunit A	11820	C	G
Rv0667	rpoB	DNA-directed RNA polymerase subunit beta	764666	G	A
Rv0682	rpsL	30S ribosomal protein S12	781687	A	G
Rv0682	rpsL	30S ribosomal protein S12	791249	C	T
Rv1267c	embR	transcriptional regulator EmbR	1416222	A	G
Rv1267c	embR	transcriptional regulator EmbR	1416232	A	G
Rv1592c	Rv1592c	hypothetical protein	1803265	G	A
Rv1694	tlyA	16S/23S rRNA (cytidine-2'-O)-methyltransferase TlyA	1924008	G	A
Rv1908c	katG	catalase-peroxidase	2155168	C	G
Rv1908c	katG	catalase-peroxidase	2158109	T	C
Rv2243	fabD	malonyl CoA-acyl carrier protein transacylase	2517129	A	G
Rv2243	fabD	malonyl CoA-acyl carrier protein transacylase	2521342	T	C
Rv2247	accD6	acetyl-/propionyl-CoA carboxylase subunit beta	2521428	A	G
Rv2247	accD6	acetyl-/propionyl-CoA carboxylase subunit beta	2522955	T	G
Rv3793	embC	arabinosyltransferase C	4243346	A	G
Rv3794	embA	arabinosyltransferase A	4243460	C	T
Rv3794	embA	arabinosyltransferase A	4248003	A	G
Rv3795	embB	arabinosyltransferase B	4250747	TG	-
Rv3919c	gid	rRNA small subunit methyltransferase G	4407927	T	G
Rv3919c	gid	rRNA small subunit methyltransferase G	4408923	C	T



Visualization of comparative bioinformatics analysis by Venn diagram (sensitive - SUSC, MDR and XDR clinical isolates)

Conclusion

➢ 20 whole genomes of *M.tuberculosis* with different drug resistance profiles were sequenced using Roche GS FLX+ platform.

➢ The structure of the *M.tuberculosis* strains genotypes determined by spoligotyping. The results of genotyping *M.tuberculosis* strains showed the prevalence of *M.tuberculosis* W-Beijing (83,3%) strains among studied clinical isolates.

➢ Four common genetic loci were found only in MDR and XDR isolates (genes PE_PGRS24, PPE24, PPE5, PE_PGRS56), that were not found in other isolates. Proteins of this family of genes may play a role of virulence factors and contribute to a successful infection.

References

1) Global Tuberculosis Report 2017. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.
2) Ministry of Health Republic of Kazakhstan. Statistic TB Review (Editor Ismailov Zh.K). National TB Center, Almaty, 2017 (in Russian).
3) Judith R. Glynn, Kristin Kremer, Martien W. Borgdorff, Mar Pujades Rodriguez, Dick van Soolingen European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis // Beijing/W Genotype *Mycobacterium tuberculosis* and Drug Resistance. Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 12, No. 5, May 2008.
4) Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 35: 907-914.

Acknowledgments

The authors thank the: - Mycobacteriology laboratory, New York State Department of Health Wadsworth Center, Albany NY, USA
- Reference Laboratory, National TB Center, Almaty, Kazakhstan.

This study was financially supported by the Ministry of Education and Science of the Republic of Kazakhstan

Ulan Kozhamkulov, PhD, National Laboratory Astana, Nazarbayev University, 53 Kabanbay Batyr ave., 010000, Astana, Republic of Kazakhstan,
Tel: +7(7172)694977, e-mail: ulan.kozhamkulov@nu.edu.kz