Характеристика мутаций штаммов Mycobacterium tuberculosis, ассоциированных с лекарственной устойчивостью на основе данных полногеномного секвенирования (CARE)

Characterization of mutations in Mycobacterium tuberculosis strains associated with drug resistance based on genome-wide sequencing data

Ekaterina Chernyaeva, Mikhail Rotkevich, Olga Lebedenko, Alla Lapidus, Dmitriy E. Polev, Natalia Solovieva, Viacheslav Zhuravlev, Piotr Yablonsky, Stephen J. O'Brien

St. Petersburg State University
St. Petersburg Research Institute of Phthisiopulmonology

9th Congress of the National Association of Phthisiologists of Russia.
November 23-24, 2020 St. Petersburg.

The aim of our study

- Compare genomic landscapes of *M.tuberculosis* isolates obtained from patients with different clinical features of the disease – pulmonary (PTB) and extrapulmonary TB (XPTB).
- Identify major mutations known to be associated with drug resistance to first-line and second-line anti-tuberculous drugs.

Materials of the study

- A total of 72 pulmonary and 73 extrapulmonary *M. tuberculosis* isolates were collected from different patients within the period from 2007 to 2014 in 40 different regions of the Russian Federation.
- 18 isolates received from extrapulmonary localization were collected from patient with generalized TB.
- 120 isolates from HIV-negative and 25 – from HIV-infected patients

- RIF
- INH
- SM
- EMB
- PZA
- ETH
- OFL
- KM, AM, CAP
- CYCLO
- PAS

Experiment

M. tuberculosis cultures

DNA extraction

Illumina MiSeq WGS sequencing

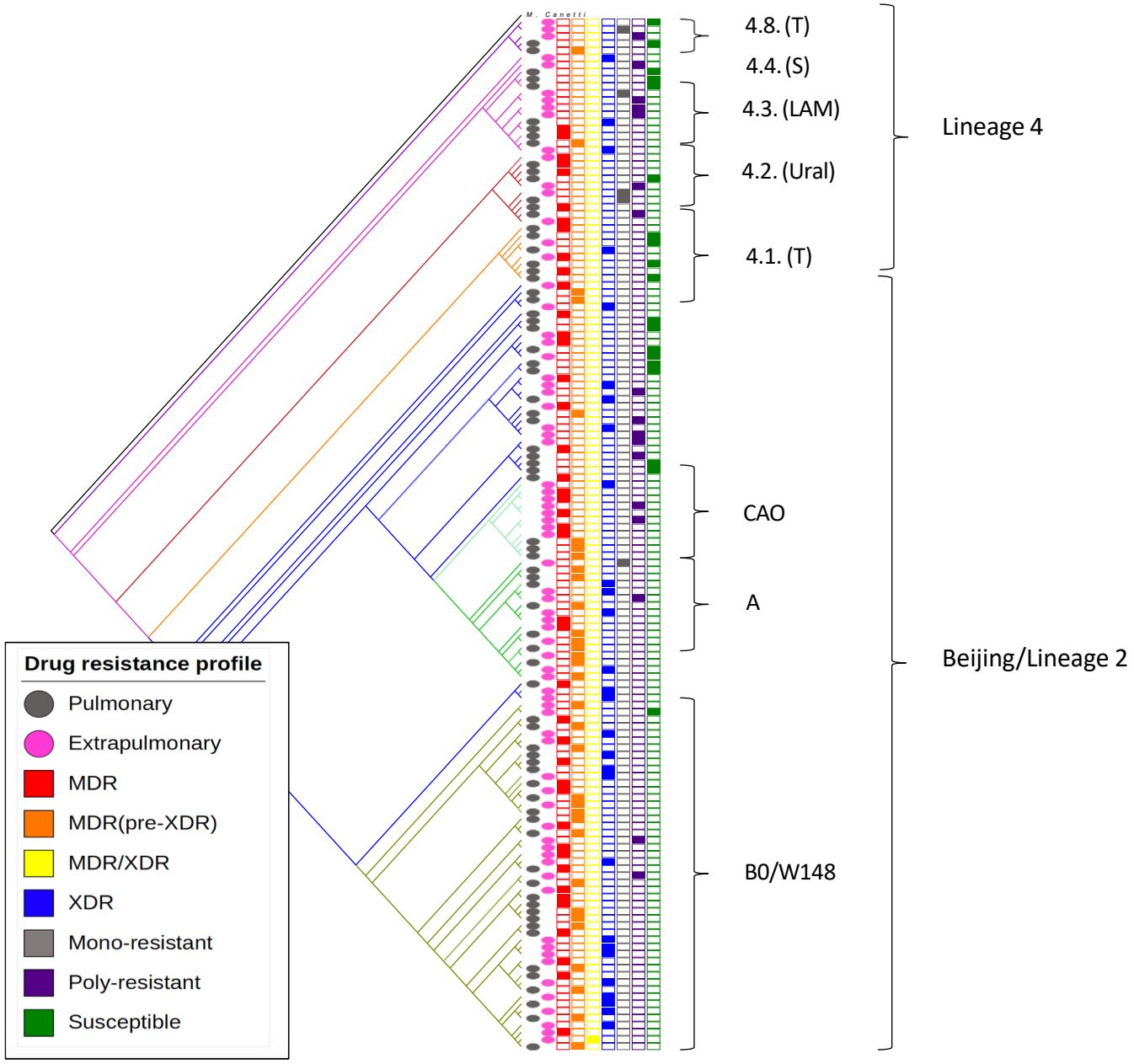
Reads alignment to H37Rv for SNV and InDels calling
(bowtie2, samtools, FreeBayes)

Phylogenetic analysis (RaxML)

in silico spoligotyping (SpoTyping)
PhyTB classification

Statistical analysis

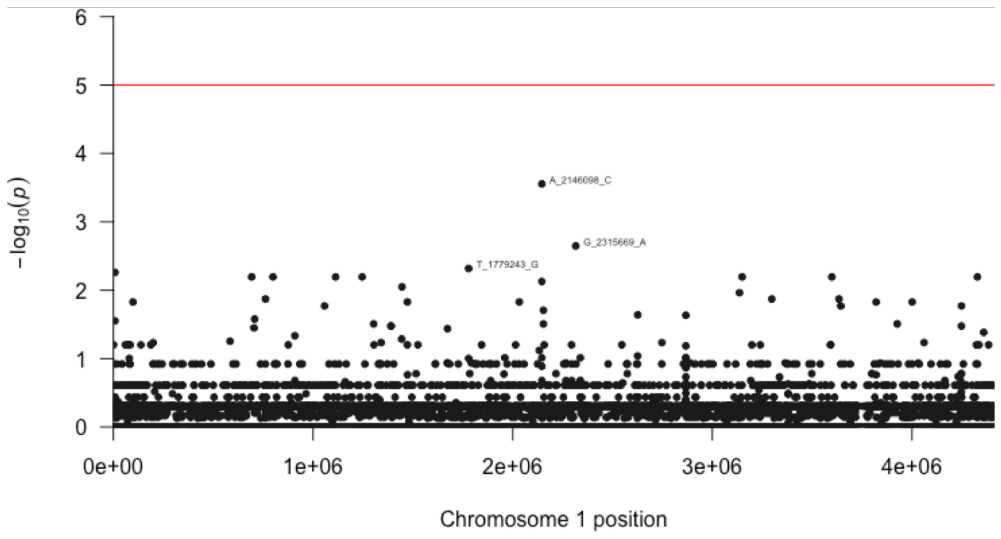
M. tuberculosis phylogenetic analysis



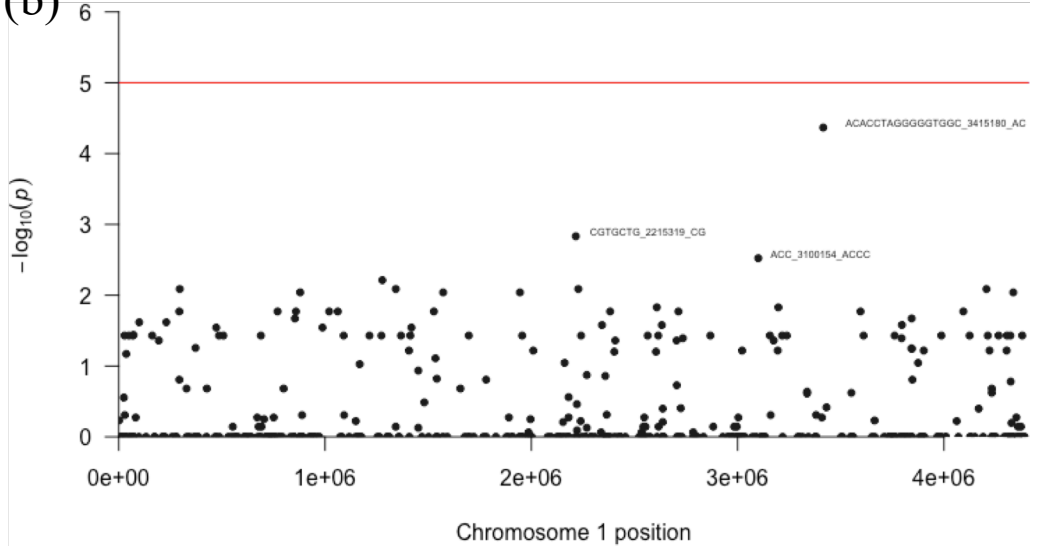
	Pulmonary TB		Extrapulmonary TB		Total	
Genetic group	N	%	N	%	N	%
Beijing:	48	66.67	60	82.19	108	74.48
Beijing unclustered	16	22.22	15	20.55	31	21.38
Beijing B0/W148	26	36.11	23	31.51	49	33.79
Beijing Clade A	3	4.17	14	19.18	17	11.72
Beijing CAO	3	4.17	8	10.96	11	7.59
4.1	5	6.94	3	4.11	8	5.52
4.2	6	8.33	4	5.48	10	6.9
4.3	11	15.27	2	2.74	13	8.97
4.4	1	1.39	0	0	1	0.69
4.8	1	1.39	4	5.48	5	3.45

Genome-wide association analysis for XPTB and PTB M. tuberculosis isolates

(a)



(b)



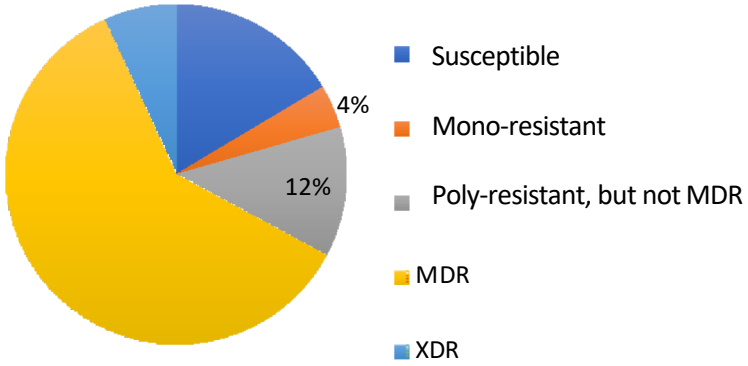
SNPs and InDels specific for phylogenetic groups (found in 95% of samples in each group) were excluded from the analysis.

There were no SNPs (a) and InDels (b) found to be associated with TB tissue localization, according to FET.

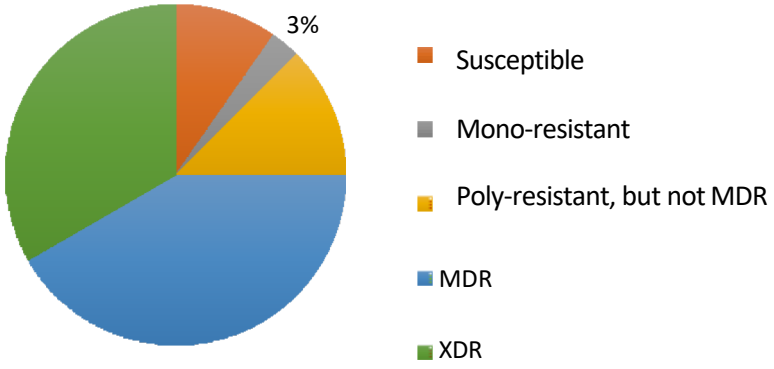
Drug resistance of PTB and XPTB strains

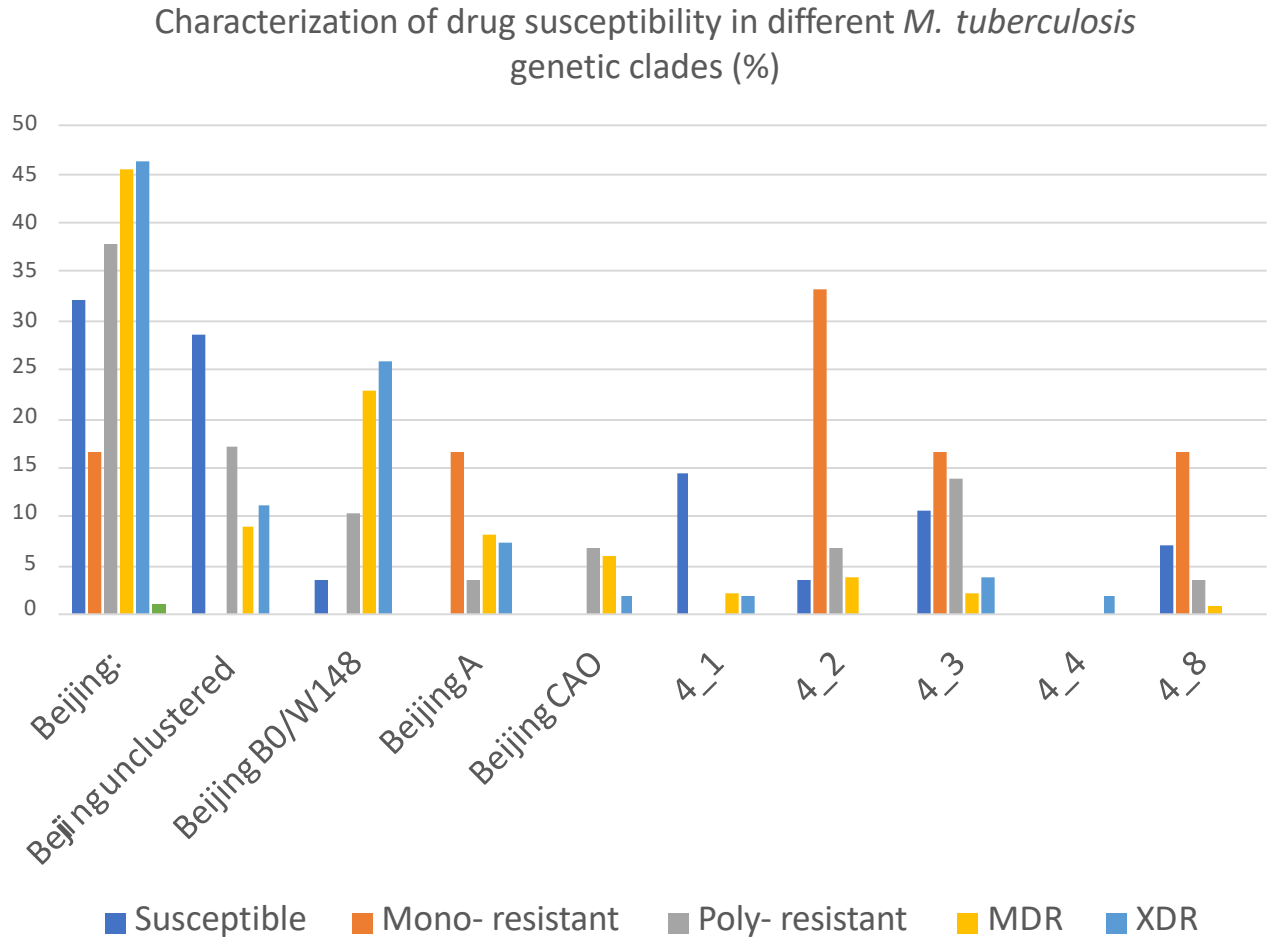
Characterization of drug susceptibility of *M. tuberculosis* from different localizations

Pulmonary



Extrapulmonary





High association of XPTB with M/XDR ($p= 6.133*10^{-05}$)

Mutations associated with drug resistance

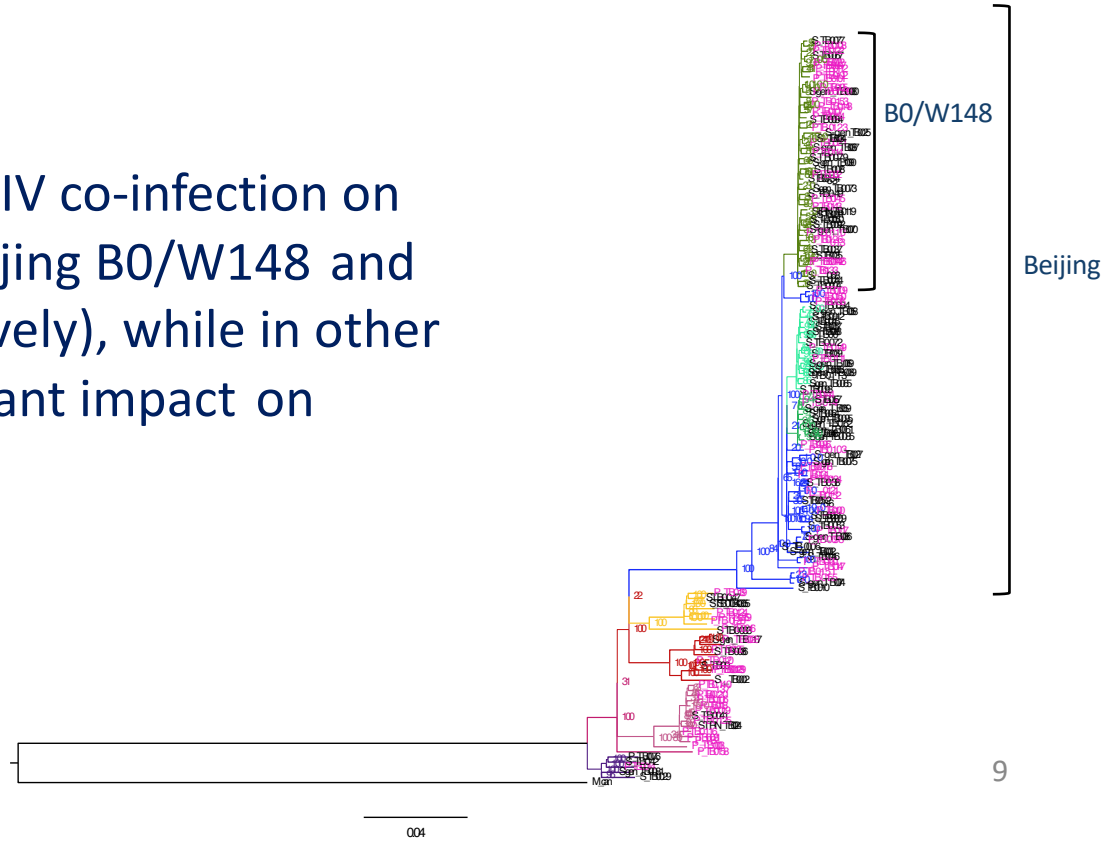
- INH-resistance has been associated with mutation S315T in *katG* gene in 95.87 % of INH-resistant strains.
- RIF-resistance was associated with mutations in *rpoB* 81-bp core region, 96.15% of resistant isolates carried mutations in this genome region.
- SM-resistance is associated with mutations in *rpsL*, *rrs* and *gid* genes in 98% of SM-resistant cases.
- Mutations in promoter region of *eis* gene were identified in 18 KM-resistant and 13 KM-susceptible isolates.
- It was previously shown that mutations in regulatory region of *whiB7* gene can indirectly influence KM-resistance, however more likely are associated with SM-resistance. We identified only two KM-resistant and 3 susceptible isolates with mutations in the region upstream *whiB7* and all of these mutants were resistant to SM.
- OFL-resistance could be explained by mutations in *gyrA* or *gyrB* quinolone resistance-determining regions in 90.48% cases.
- Mutations in *embB* gene between codones 296 and 497 were detected in 63 resistant and 27 susceptible to EMB isolates.
- Mutations in *embC-embA* intergenic region (8-16 nucleotides upstream *embA* gene) were detected in 16 EMB-resistant and 7 EMB-susceptible isolates.
- Point mutations were detected in *pncA* and *rpsA* genes in 31 PZA-resistant isolates (67.39%)

HIV co-infection

Diagnosis	HIV+	HIV-	N
PTB	3	69	72
EPTB	14	41	55
Generalized TB	8	10	18

HIV-infection increases the probability of TB generalization or development of active extrapulmonary TB ($p= 6.579*10^{-05}$ and $p=0.000998$)

Statistical analysis allowed to detect the influence of HIV co-infection on extrapulmonary TB development in genetic groups Beijing B0/W148 and unclustered Beijing ($p= 0.0013$ and $p= 0.0068$ respectively), while in other genetic groups HIV co-infection did not make a significant impact on extrapulmonary TB development.



Conclusions

- XPTB *M.tuberculosis* isolates belong to Beijing genetic group (82.2%) more often than PTB (66.7%). High prevalence of Beijing strains among extrapulmonary *M. tuberculosis* samples can explain a very high association of Extrapulmonary TB with M/XDR identified in our study.
- Analysis of mutations, associated with bacterial resistance to first- and second-line TB drugs allowed to identify that relatively high proportion of INH-, RIF-, SM- and OFL- resistant isolates had standard SNVs predictive of drug-resistance.
- Mutations in *pncA* and *rpsA* genes that were known to be involved in development of resistance to PZA were identified. Most of PZA-resistant isolates had mutations in *pncA* gene mentioned in previously published catalog of mutations associated with resistance to PZA, 4 PZA-resistant isolates carried mutations in *rpsA* gene. However, we detected few mutations, that were mentioned as associated with resistance to PZA and among several PZA-susceptible isolates.
- Detection of mutations associated with drug resistance among susceptible isolates may indicate the presence of low number of drug resistant clones in *M.tuberculosis* population and might be a signal for correction of TB therapy in case it was developed based on phenotypic data.
- HIV-infection increases the probability of TB generalization or development of active extrapulmonary TB.

convPhy – a new tool for identification genomic markers associated with drug resistance

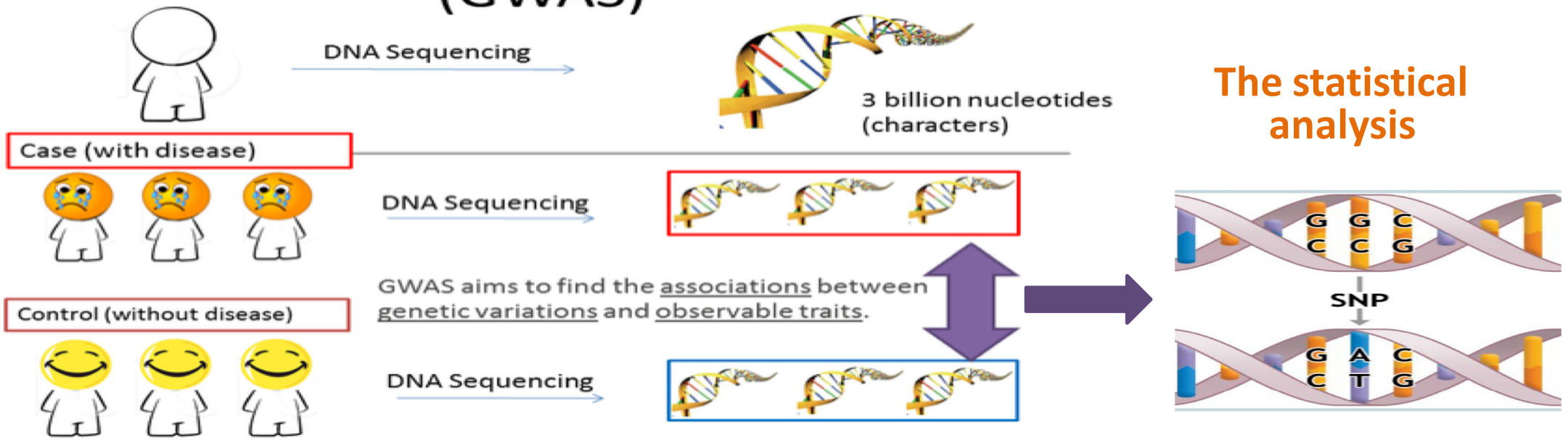
ConvPhy - phylogenetic convergent test

GWAS in general

Basic anatomy of GWAS:

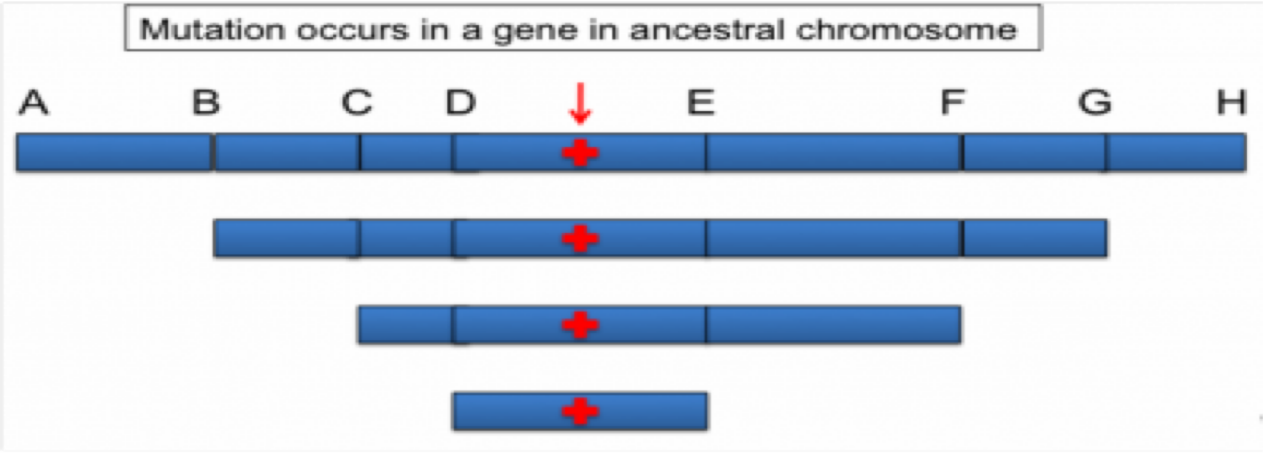
- **Count alleles for each polymorphic site**
- **Correct for multiple comparisons**

- **Standard**
- **With Incorporating interactions**
- **Correction for multiple testing**



Confounders in GWAS

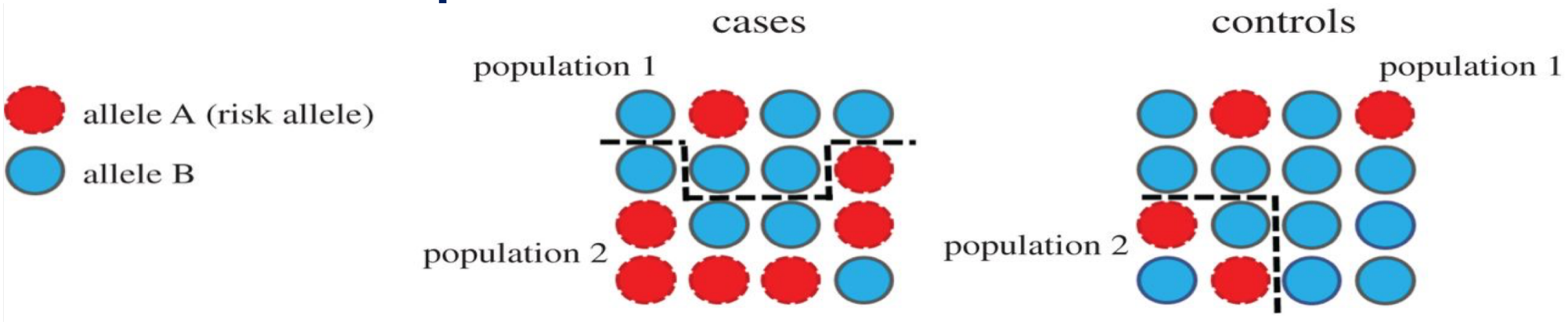
LD linkage disequilibrium



Multiple testing correction

Multiple testing correction for non-independent, interconnected gene sets

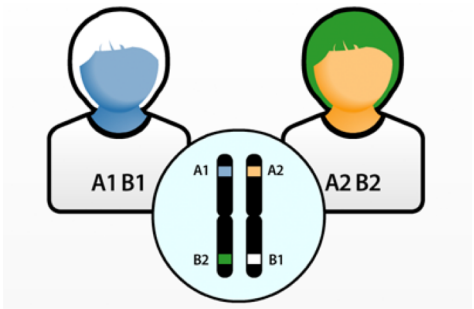
Population stratification

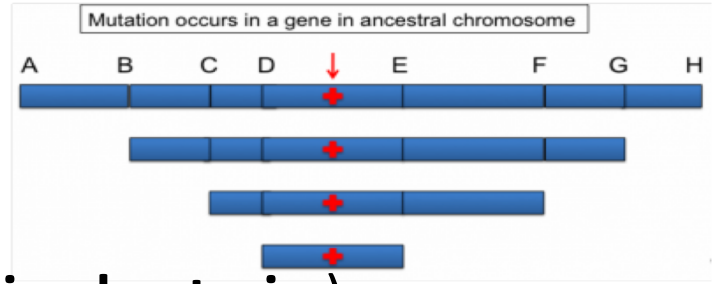


Bacteria have more strong LD and population stratification

	Eukaryote	Bacteria
Ploidy	diploid	haploid
Genetic re-assortment	**sexual** 1 homologous recombination 2 chromosome segregation	**asexual** 1 short gene conversion 2 horizontal gene transfer

disruption of LD





- clonal frame (frequently recombining bacteria)
- clonal structure (clonal bacteria)

The separation of causative variants from passive linked loci is potentially a difficult problem

Natural selection vs genetic drift

Human

Phenotypes tend to be shaped by genetic drift

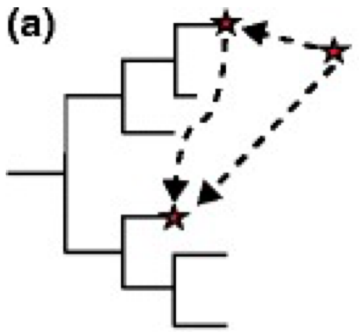
Bacteria

Phenotypes tend to be shaped by strong natural selection

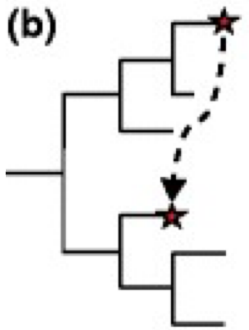
relatively small samples of bacterial genomes should be sufficient to identify causal mutations

15

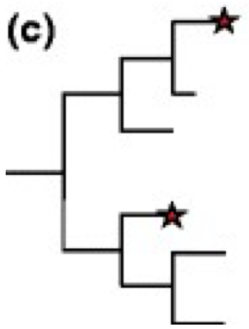
by Shapiro J., Chen P. The advent of genome-wide association studies for bacteria // Current opinion in microbiology. 2015. Vol. 25. P. 17-24.



Horizontal gene transfer: the movement of mobile genetic elements, via conjugation, transformation or transduction from one lineage to another, or the acquisition of genetic elements by distinct lineages independently from a common donor.



Recombination: DNA from a distant relative enters a cell via transduction, conjugation or transformation, and by homologous recombination replaces the existing DNA.



Recurrent mutation: the same mutation evolves independently in distinct lineages.

Natural mechanisms for introducing homoplasious mutations into the genomes of bacterial populations

for clonal bacteria like *M.tuberculosis*

by Timothy D.R., Ruth C. M. Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: a new direction for bacteriology // Genome Medicine. 2014. V. 6. N. 109. P. 1-11.

ConvPhy. Preparing data

Indian *M.tuberculosis* isolate

N = 223

SRA: PRJNA235851

Russian *M.tuberculosis* isolate

N = 137

SRA: PRJNA352769

WGS

Raw reads

Alignment to reference genome H37rv

SNP calling

Filtering SNP

Annotation SNP

Reconstruction phylogenetic tree.

Isoniazid

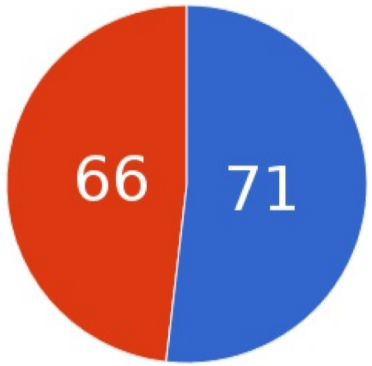
Rifampicin

Ethambutol

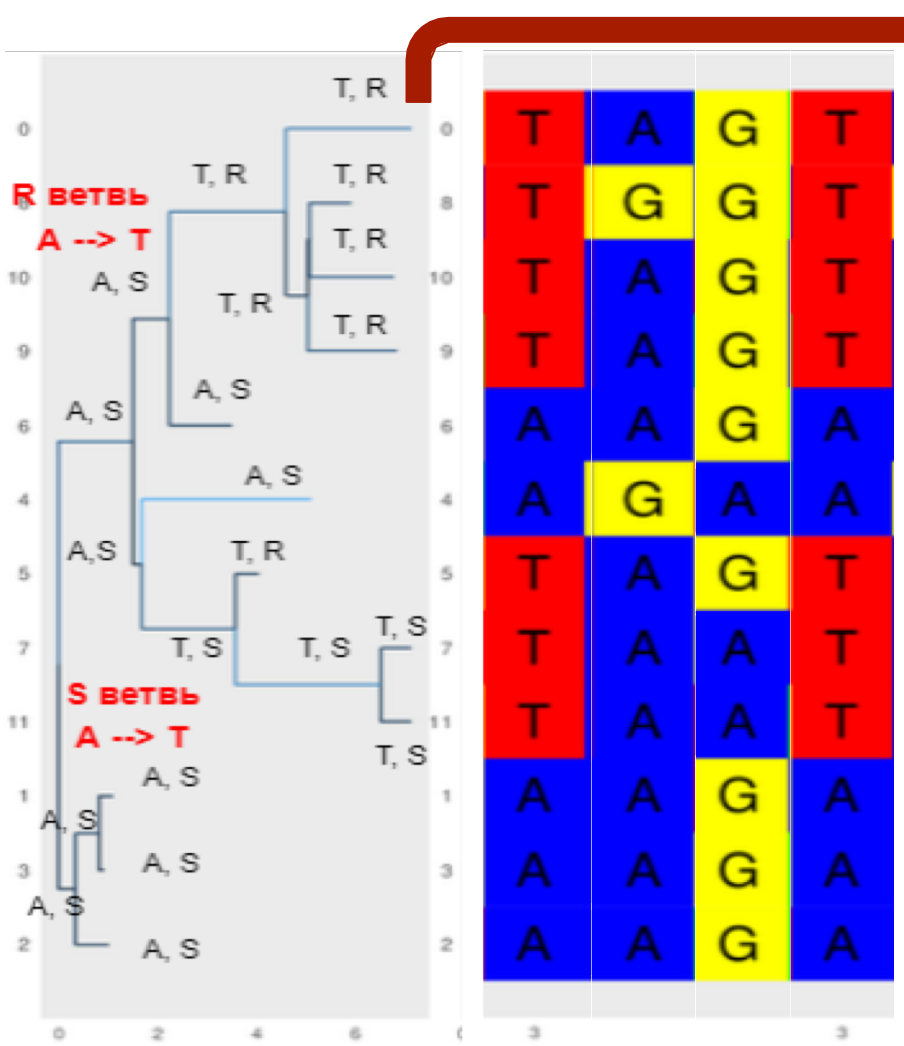
Extrapulmonary

Pulmonary

localization

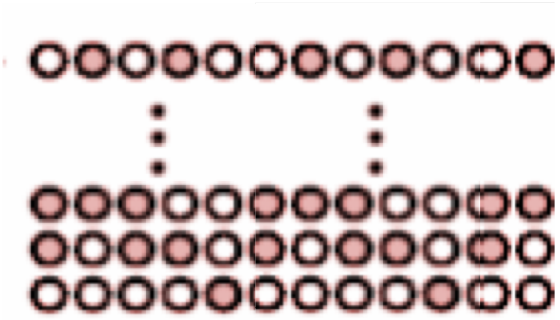


convPhy pipeline



R1, S1 - for every SNP

Permutation test
10 000 iterations:



pvalue - fraction of samples with
R > = R1, S >= S1

Input data

1) Phylogenetic tree

2) List of SNPs

3) Binary phenotype:

R - positive (resistant)

S - negative (sensitive)

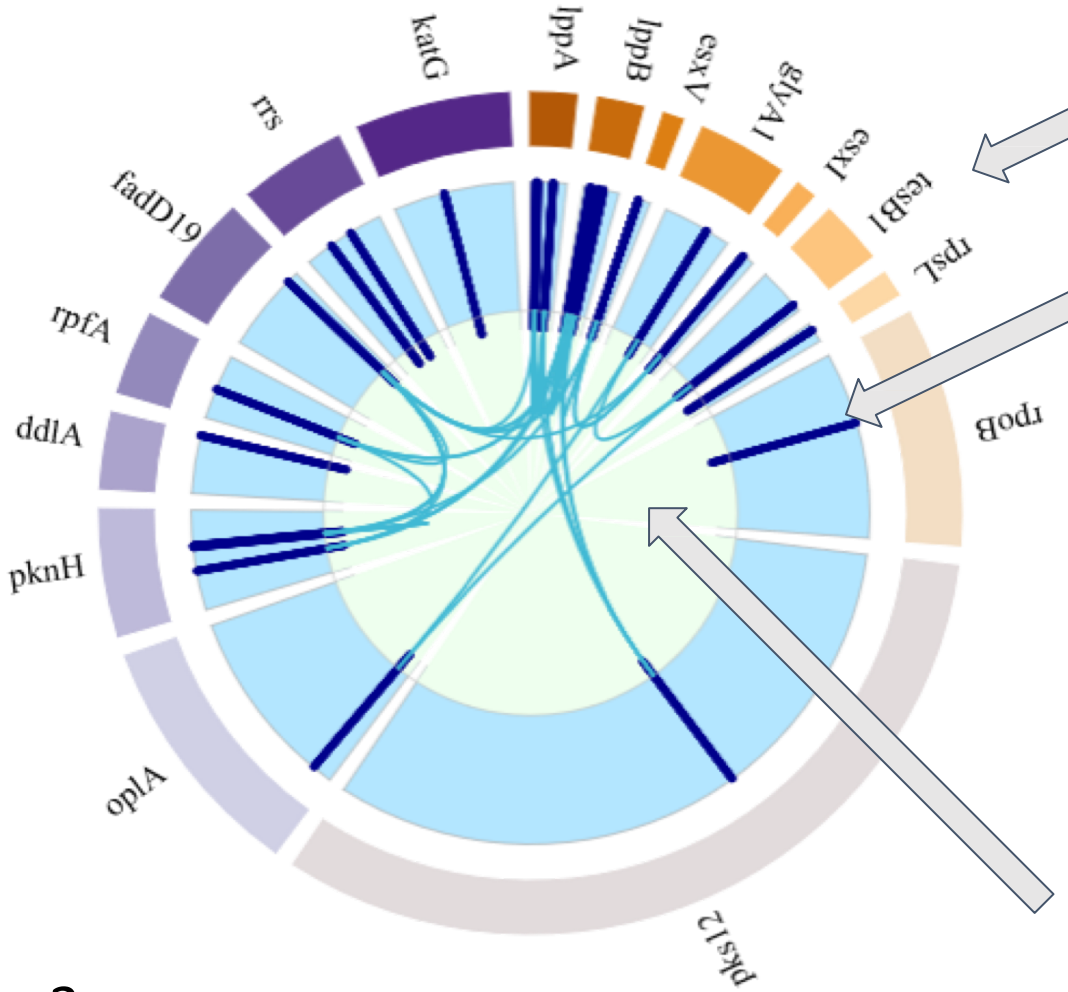
Output data

annotated significant SNPs

Visualization output of convPhy

Results

convPhy



Outside circle: genes

SNPs

Cooperation interaction

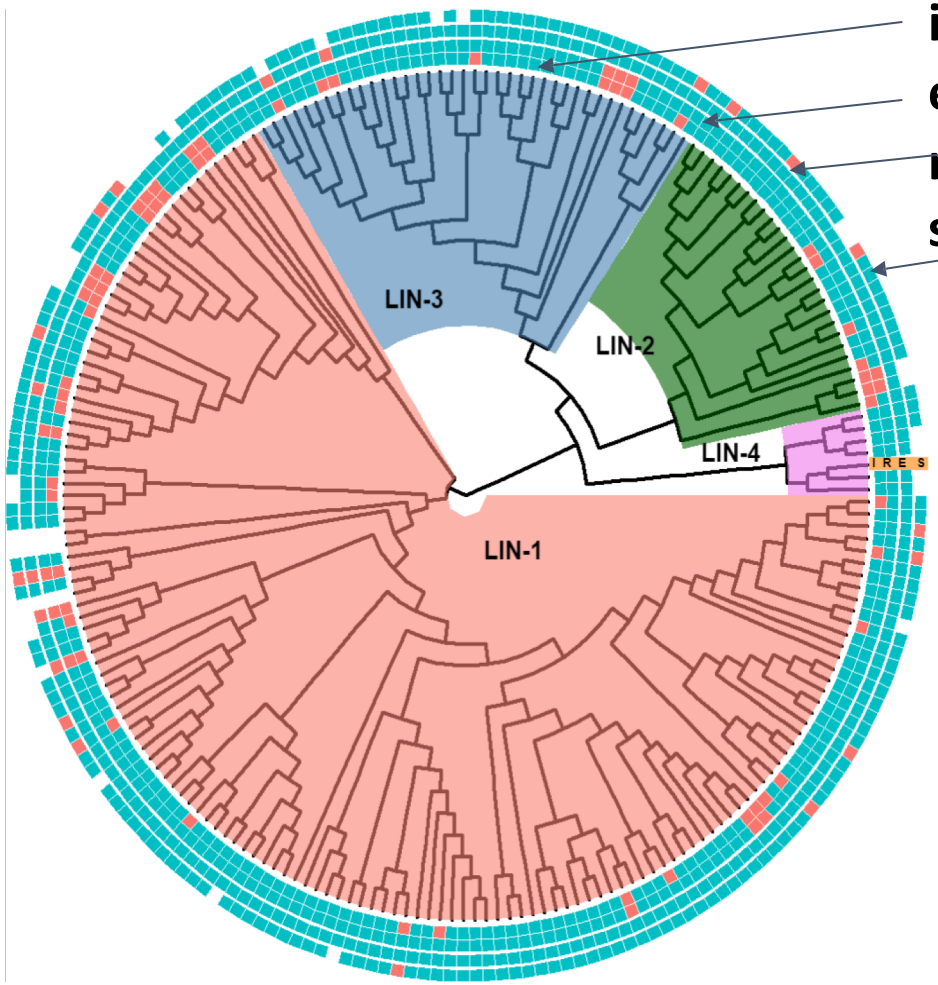
t-test

	alt pos1	ref pos1
alt pos2		
ref pos2		

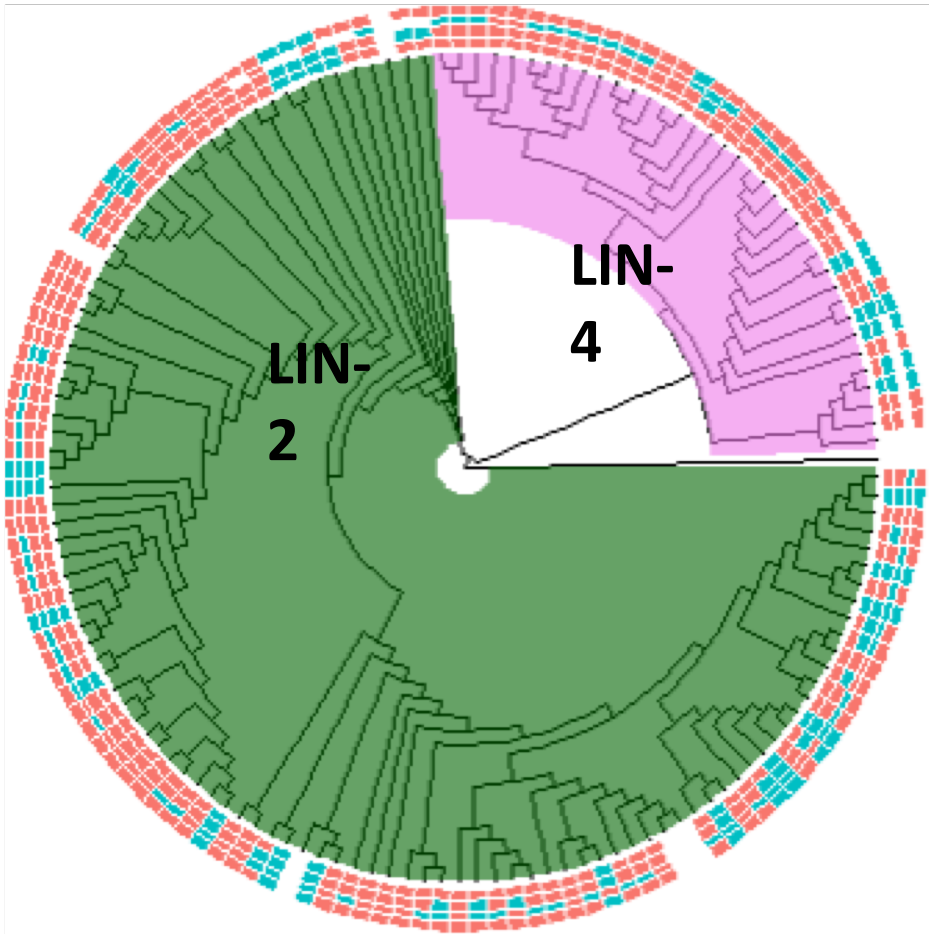
source code python3:
https://github.com/OOLebedenko/convPhy/
Visualization with R library BioCircos

Phylogenetic structure

Indian population

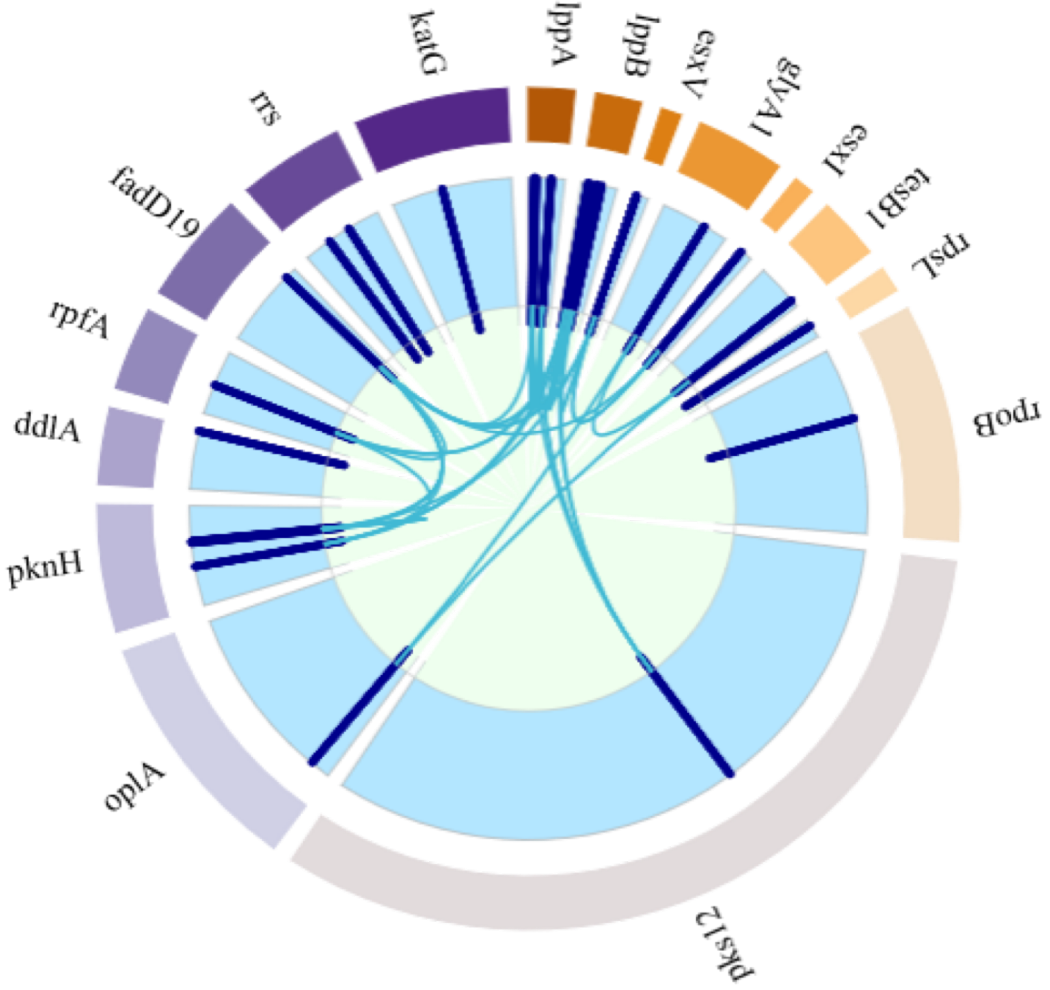


Russian population



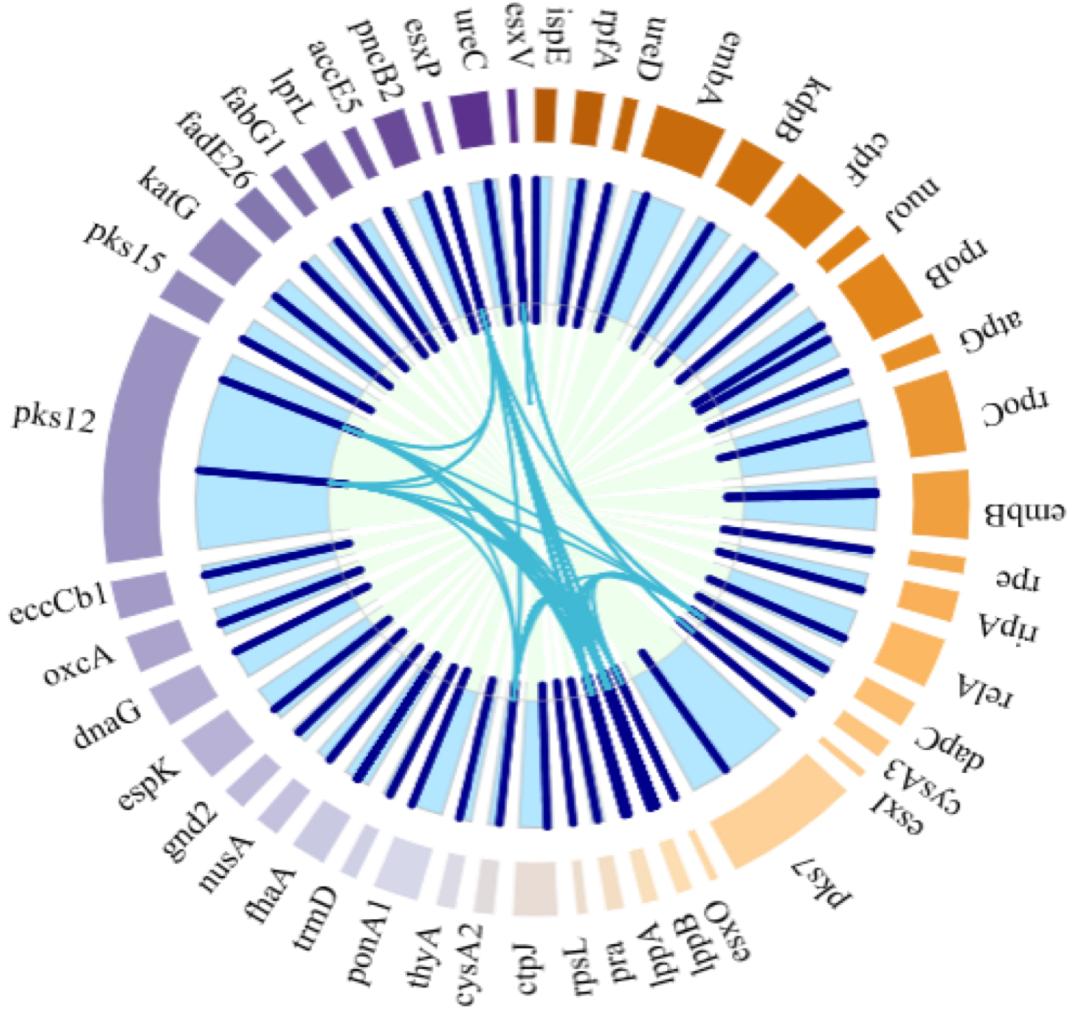
Cooperation intergarion. Resistance to isoniazid

Indian population



S315T *katG* 68% isolates

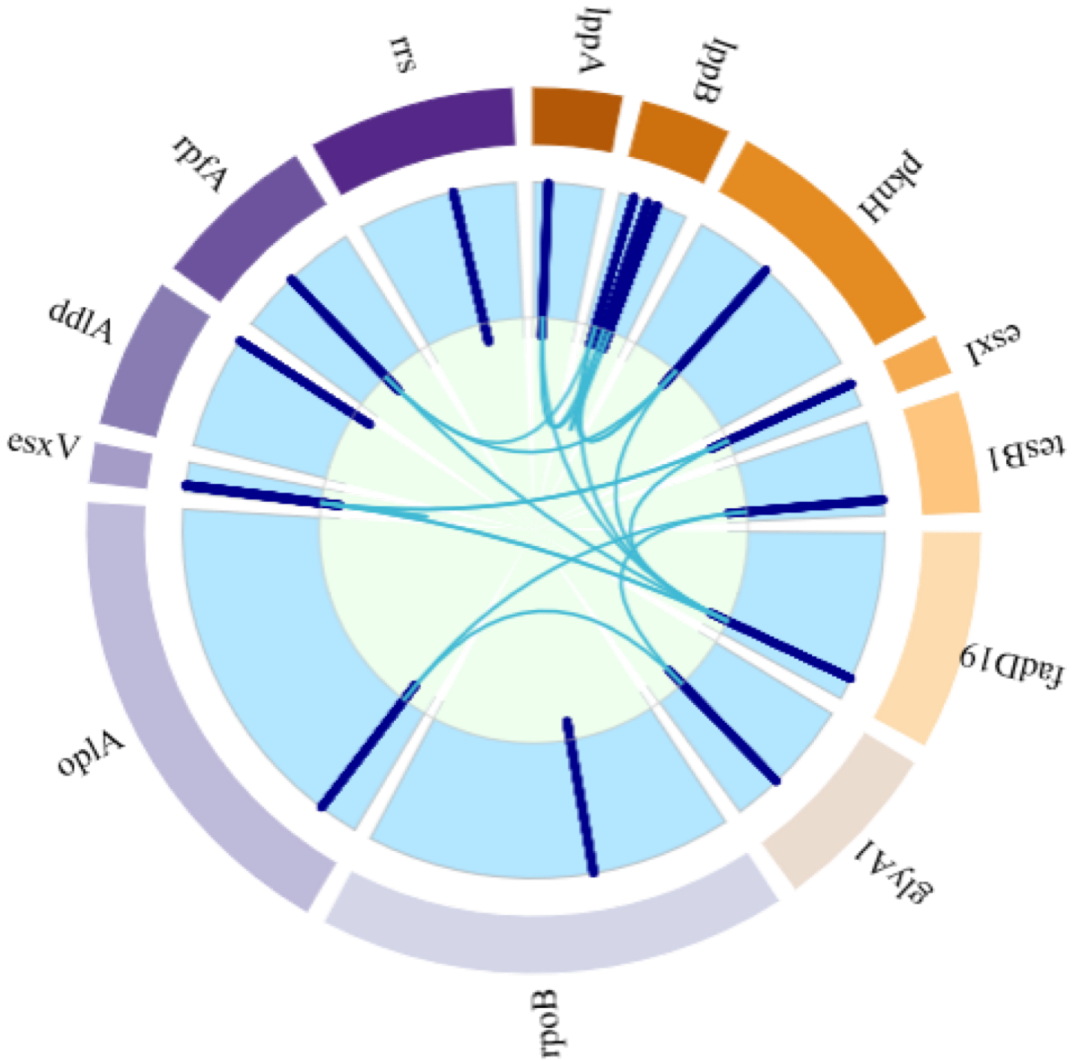
Russian population



S315T *katG* 100% isolates

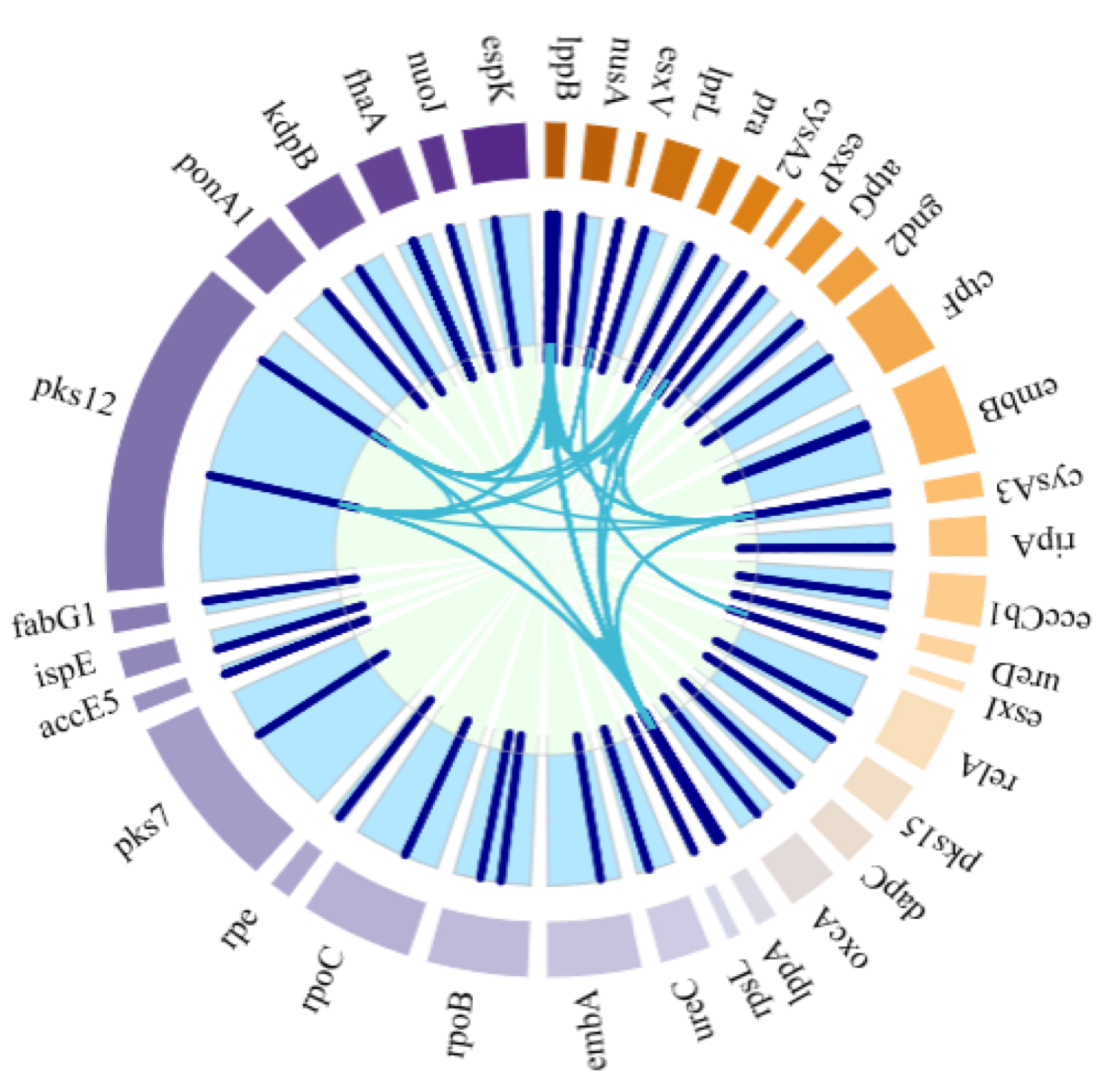
Cooperation interation. Resistance to rifampicin

Indian population



S450L *rpoB* 40% isolates

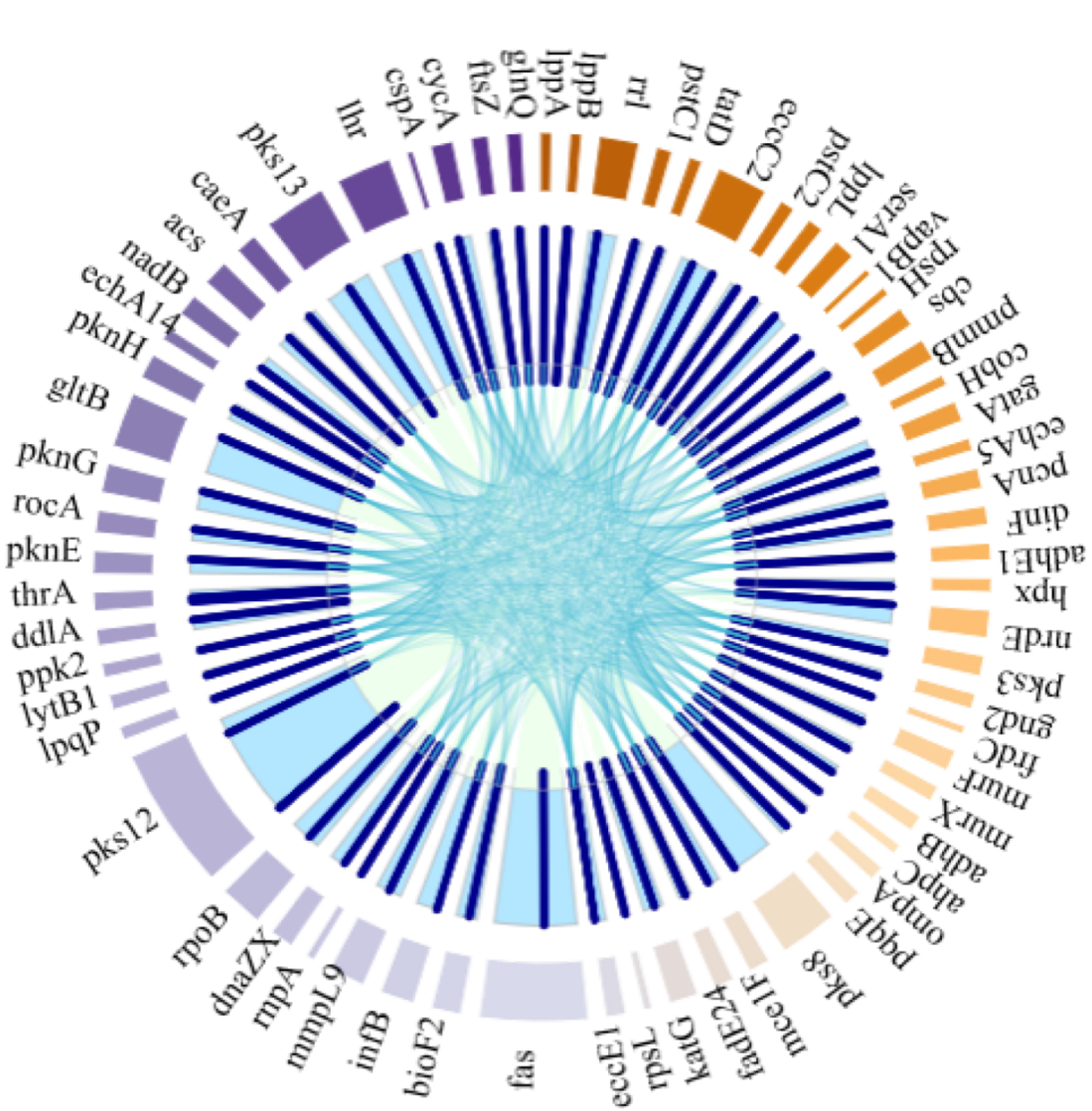
Russian population



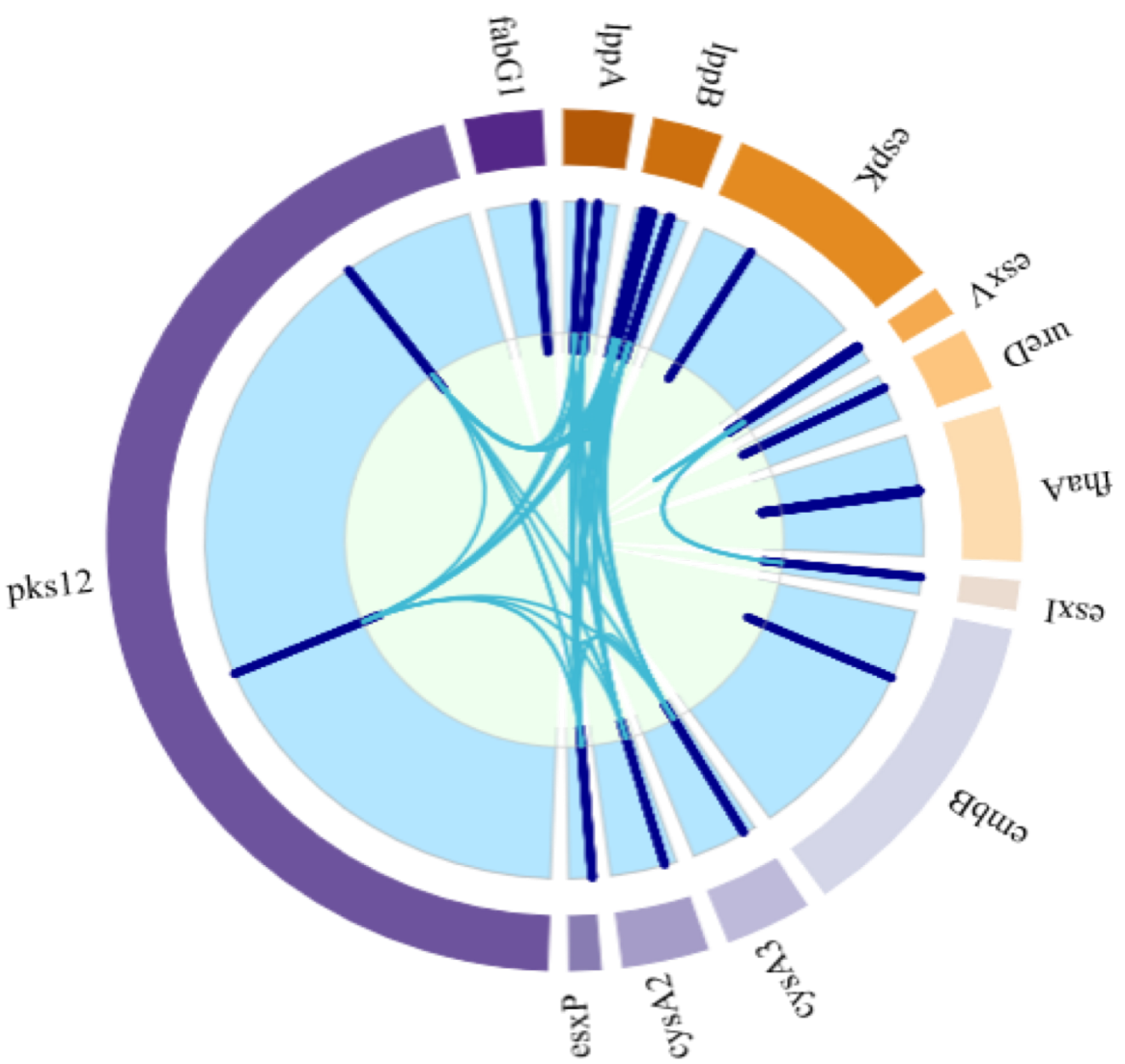
S450L *rpoB* 80% isolates

Cooperation interation. Resistance to ethambutol

Indian population



Russian population



Conclusions

1.ConPhy demonstrated high level of concordance with associations described in Indian validation SNPs set.

2.In Russian isolates convPhy revealed association pathogen localization with mutations in PE/PPE genes, presumably participating in virulence.

3.conPhy detected mutation previously not associated with drug resistance in esxV, pks12, lppa, lppb, pknH, fadD19, glyA1, rpfa genes and cooperative interaction of single SNPs in isolates M. tuberculosis whose resistance to isoniazid and rifampicin are not explained by known mutations S315T katG and S450L rpoB.

Funding:

The authors acknowledge the contribution of CARE Consortium funded by the European Union's Horizon 2020 programme and the Ministry of Science and Higher Education of the Russian Federation (unique project identifier RFMEFI61019X0020)

Acknowledgments:

Scientific research was performed using equipment of the Resource Center «Biobank», Research park of St. Petersburg State University.

Ekaterina Chernyaeva, echernya@gmail.com