doi 10.18699/bgrs2024-5.2-14

Genomic technologies in the study of the phylogeny of plague pathogen and certification of foci of the Caspian region and Central Asia

Balykova A.N.*, Kovrizhnikov A.V., Eroshenko G.A., Krasnov Ya.M., Kutyrev V.V.
Russian Research Anti-Plague Institute "Microbe", Federal Service for Surveillance in the Sphere of Cosumers Rights Protection and Human Welfare, Saratov, Russia
** alinabalnik@gmail.com*

Key words: *Yersinia pestis*; strains of medieval biovar; molecular-genetic analysis; evolution; phylogenetic analysis

Motivation and Aim: The importance of studying the genomic portrait and evolution of the causative agent of natural foci of a particularly dangerous infectious disease - plague, which left an indelible mark in history and still retains its pandemic potential, is beyond doubt. In this work, *Yersinia pestis* strains from plague foci of the Caspian region and adjacent foci of Central Asia, which are located in four countries: Russia, Kazakhstan, Turkmenistan, and Uzbekistan, were studied. The medieval biovar of the main subspecies of the plague pathogen, *Y. pestis subsp. pestis*, circulates in these foci. This biovar is widespread in 33 out of 45 existing plague foci, which is 93.3% of the total area of foci in the CIS countries. Its strains are highly virulent and epidemically significant [1]. The reasons for the high adaptive characteristics of the medieval biovar and its rapid spread in the twentieth century have not yet been determined. We have previously established that at the beginning of the twentieth century two branches of the medieval biovar, 2.MED1 and 2.MED4, which strains were etiologic agents of plague outbreaks with high mortality in the period 1912–1950, were widespread in the territory of the foci of the Northern Caspian region [2]. The aim of this study was to conduct comprehensive MLVA25-/SNP-/CRISPR typing and to estimate the genotype diversity of *Y. pestis* strains in the foci of the Caspian region and Central Asia.
Methods and Algorithms: 350 strains of *Y. pestis* isolated in 1912–2015 from carriers, vectors and humans in 25 natural plague foci of the Caspian Sea, Caucasus and Central Asia were studied. Whole genome sequencing of *Y. pestis* strains was performed using MGI (DNBSEQ-G50RS) platform with MGIEasy FS DNA Library Prep Set and MGIEasy UDB PF Adapter Kit A according to manufacturer's instruction. Fastp v0.23.3 and Trycycler v0.5.4 were applied for data processing. Fragment sequencing was performed using the ABI PRISM 3500XL platform (Applied Biosystems, USA) with 3500 Series Data Collection Software for data processing. For comparative SNP/VNTR/CRISPR analysis developed author's bioinformatics programs (VNTRfinder v0.4, SNPgenotyper v0.1) were used. The Hunter-Gaston diversity index (D) was used to assess the discriminatory power of the typing method.
Results: According to the results of molecular genetic analysis and phylogenetic reconstruction of 81 strains of *Y. pestis* on the basis of 2360 core SNPs, the population structure and genetic profile of the main populations of *Y. pestis* from 25 plague foci of the Caspian region and Central Asia for the period 1912–2015 were determined. WG-SNP analysis of the core genome of 740 strains of the major main and non-main

subspecies of *Y. pestis* was performed to determine stable SNP markers. Typing was performed using the developed programs SNPgenotyper v0.1 and SimGA v.0.4. 59 polymorphic nucleotides (SNPs) were determined, which determine 23 genovariants spread in different epidemic periods of activity of the studied plague foci (Fig 1.).

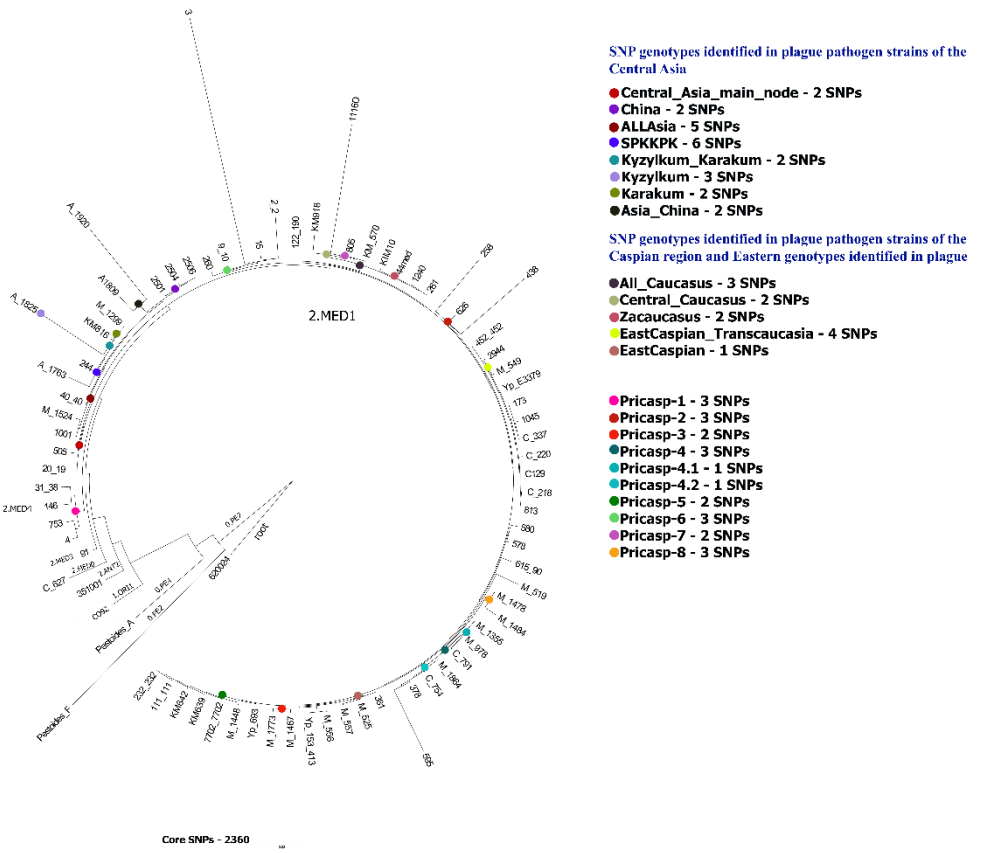

Fig. 1. Population structure of the medieval biovar of *Y. pestis* and established SNP-genotypes of strains from plague foci of the Caspian region and Central Asia of the period 1912–2015.

The Hunter-Gaston discriminating index for this typing scheme was D = 0.9, which ensures the development of SNP panels for the differentiation of populations of the medieval biovar of the plague pathogen. A method of identification of detected SNPs in the plague foci of the Northern and Northwestern Caspian region by fragment sequencing was also developed.

VNTR loci were searched using the MLVA25 scheme and CRISPR loci using the VNTRFinder v0.4 program. Based on the data received VNTR loci (Variable number tandem repeats) with high resolution for differentiating these populations were identified. The loci ms05, ms06, ms07, ms15, ms46, ms56, ms62, ms70, ms74 (number of locus alleles > 3, allele polymorphism index h > 0.5) had the highest variability. Comparative analysis of the locus composition of CRISPR elements, which are

responsible for the immune system of bacteria and can store information about the past interventions of phages, was carried out for the loci characteristic of the plague pathogen, YPa (YP1), YPb (YP2), and YPc (YP3) [3]. The vast majority had YPa (YP1) size of 305 bp, YPb (YP2) locus had a size of 321 bp, and YPc (YP3) locus had a size of 321 bp. The CRISPR loci possessed the following spacer profile: YPa a1-a2-a3; YPb b1-b2-b3-b3-b4`; YPc c1-c2-c3. This profile was identified previously as characteristic of the medieval biovar of the main subspecies [4]. Some strains had a unique spacer profile, which can be further used as a marker in genetic certification of foci of the plague pathogen.

Conclusion: Based on the results of a comprehensive analysis of 350 strains of *Y. pestis* isolated in 1912–2015 in 25 natural plague foci in four countries: Russia, Kazakhstan, Turkmenistan and Uzbekistan, new data on the evolution and genetic variability of populations of the highly virulent and epidemically significant medieval biovar of the plague pathogen were obtained. SNP-/VNTR-/CRISPR loci of the genome of the medieval biovar were analyzed using the author's computer programs. 59 SNPs were identified and SNP panels are being developed for molecular genetic certification of plague foci. The profile of CRISPR elements and spacers (YPa a1-a2-a3; YPb b1-b2-b3-b4`; YPc c1-c2-c3) was established; VNTR loci providing differentiation of genetic populations of *Y. pestis* medieval biovar of different origin were determined. The obtained results are important for the study of variability of the genome architecture of the plague pathogen and for improving the efficiency of molecular epidemiologic tracing of the origin of pathogenic strains during molecular epidemiologic monitoring of plague foci.

Funding: This work did not receive any financial support from commercial or non-commercial organizations.

References

1. Kutyrev V.V. et al. Phylogeny and Classification of *Yersinia pestis* Through the Lens of Strains From the Plague Foci of Commonwealth of Independent States. *Front. Microbiol*. 2018;9:1106. doi 10.3389/fmicb.2018.01106
2. Eroshenko G.A. et al. Evolution and circulation of *Yersinia pestis* in the Northern Caspian and *Northern* Aral Sea regions in the 20th–21st centuries. *PLoS One*. 2021 Feb 11;16(2):e0244615. doi 10.1371/journal.pone.0244615
3. Pourcel C., Salvignol G., Vergnaud G. Pourcel C. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*. 2005;151(Pt 3):653-663. doi 10.1099/mic.0.27437-0
4. Cui Y. et al. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced palindromic repeats. *PLoS One*. 2008;3(7):e2652. doi 10.1371/journal.pone.0002652