Evolution of *Burkholderia* spp

Olga O. Bochkareva, Elena V. Moroz, Iakov I. Davydov

IITP RAS, B. karetny per. 19, Moscow Russia, olga.bochkaryova@gmail.com

We analyzed 127 complete genomes from the genus *Burkholderia*. The genus is ecologically diverse; for example, *B. mallei* and *B. pseudomallei* are pathogens causing glanders and melioidosis, respectively, in human and animals; *B. glumae* is a pathogen of rice; *B. xenovorans* is an effective degrader of polychlorinated biphenyl, used for biodegradation of pollutants; *B. phytofirmans* is a plant-beneficial endophyte that may trigger disease resistance in the host plant.

In bacteria with multiple chromosomes, the majority of genes necessary for the basic life processes usually are located in single (primary) chromosome. Secondary chromosomes contain few essential genes and are comprised of mainly niche-specific genes. Usually genes on a secondary chromosome evolve faster than genes on a primary chromosome.

Genome rearrangements such as duplications, deletions and inversions also play important roles in the evolution, as they alter the chromosome organization and gene expression in ways impossible through point mutations. Reconstruction of the history of genome rearrangements provided a base for a new type of phylogeny reconstruction algorithms.

Genomes of *Bukhloderia* consist of two or three chromosomes. The majority of single-copy common genes belong to the first chromosome, ten-fold less genes are in the second chromosome, and they are almost absent in the third chromosome, the only exception resulting from a large translocation from the first to the third chromosome in *B. cenocepacia* AU 1054.

The pan-genome analysis revealed that the pan-genome is open and saturation will be reached between 86,000 and 88,000 genes. Basic tree and the gene content tree show some differences, caused by evolution history enriched of gene gains and losses For example, reflecting niche specialization in the *B. pseudomallei* branch and genome reduction in *B. mallei*.

Reconstruction of rearrangements revealed inversions, translocations, deletion and insertions including parallel events. The reconstructed common ancestor of *Burkholderia* had 965 universal single-copy genes on the first chromosome, and 81 genes on the second chromosome.

Analysis of *B. pseudomallei* strains revealed that their chromosomes are largely collinear except for several inversions. The same inversion was found in the second chromosomes of strains that are placed on distant branches on phylogenetic tree suggesting that this event may had occurred independently. Boundaries of this inversion lie within genes encoding from the RND efflux system that are components of multidrug resistant complex.

In comparison to *B. pseudomallei*, *B. mallei* genomes harbor numerous IS elements that most likely have mediated the higher rate of rearrangements such as deletions and inversions. Inter-replichore inversions, that is, inversions with one endpoint in each replichore, are overrepresented. Moreover, within-replichore inversions turned out to be shorter than 15% of the replichore length except from two events in *B. mallei* FMH23344 that mostly overlap each other and is likely to be explained by one translocation. In contrast, the lengths of inter-replichore inversions have a much wider distribution up to replichore size. This strong selection on within-replichore inversion length caused by avoiding changes in gene position relative to lagging/leading chain if gene is placed on inverted sequence. Numerous parallel inversions were found that is likely to be explained by active recombination coupled with limited number of repeats in genomes.

Two cassettes that are spreading horizontally in. *B. cepacia* group were found. One of them is comprised of five genes that belong to iron uptake metabolic pathway, the second one is linked with lactam utilization.

This is joint work with Mikhail Gelfand. This research was supported by the Russian Foundation for Basic Research under grant 16-54-21004.