Avoidance of restriction-modification systems sites in the metagenome of Antarctic Deep Lake shows interaction of species in the metagenomic community

Marina Khachaturyan[1], Anna Ershova[1234], Ivan Rusinov[1], Anna Karyagina[123], Sergei Spirin[15], Andrei Alexeevski[15]

[1]*Lomonosov Moscow State University, Moscow 119992, Russia;*[2]*Gamaleya center for Epidemiology and Microbiology, Ministry of Health of Russian Federation, Moscow 123098, Russia;* [3]*Institute of Agricultural Biotechnology, the Russian Academy of Sciences, Moscow 127550, Russia;* [4]*Moscow Institute of Physics and Technology, the Ministry of Education and Science of the Russian Federation, Dolgoprudny, Moscow Region, 141700, Russia,* [5]*Scientific Research Institute for System Studies, the Russian Academy of Science (NIISI RAS), Moscow 117281, Russia,* asershova@gmail.com

Understanding of relationships between prokaryotes and their viruses in enviromental conditions is one of important problem in ecology and evolution of microbial communities. Prokaryotes (bacteria and archaea) use different defense systems against viruses, in particular, restriction-modification (R-M) systems. R-M systems consit of two enzymes: DNA methyltransferase, which recognizes specific sequence on DNA (recognition site) and methylate it and restriction endonuclease, which cleaves DNA if corresponding site is unmethylated. Type II R-M systems consisting of two separate restriction endonuclease and methyltransferase enzymes (orthodoxal) cause avoidance of their DNA-recognition sites in viral and prokaryotic genomes [1].

The aim of the study was to investigate if the avoidance of sites in prokaryotes and their viruses can occur as the sign of their interaction in environmental community. We analyzed metagenomes and whole genomes from antarctic Deep Lake community.

The metagenome taken from the Antarctic Deep Lake is well-characterized [2, 3], but R-M systems in this metagenome were not studied. Whole genomes of four cultivated *Haloarchaea* from Deep Lake, which form more than 70% of the community, were also used. Based on sequence similarity, 24 dsDNA and 2 ssDNA viruses were revealed in the metagenome. The archea-viral interactions according to analysis of CRISPR cassetes were taken from work Tschitschko B. et co-authors [3]. DB REBASE was used to predict R-M

system proteins and their DNA-recognition sites in metagenome. Karlin's method was used to calculate the representation of sites [4].

All types of R-M systems were detected in the metagenome. We found 320 potential R-M systems which encode 53 different DNA-recognition 4-8 length sites (Type II). Four most spread archaea genomes encode R-M system recognizing 9 different sites and each particular genome has it's own set of encoded R-M systems.

In the four archaea genomes we found 13 cases of underrepresentations for 7 sites. It is interesting that genomes of related species encode different R-M systems, but have similar site avoidance profiles which correspond to sites of R-M systems presented in the population. Thus, 8 out of 13 site avoidance cases correspond to the sites of R-M systems, which are not encoded in the particular genome, but they are in some others ones. These results testify about the horizontal transfer of R-M systems in the metagenome even between not relative organisms.

Contigs belonging to 13 different viruses were found in the metagenome. Using previous studies [3] 24 different Haloviruses can be considered as viruses detected in the community. By the site avoidance profile viruses were divided in four groups according to their site avoidance strategies and host specificity.

Site avoidance allows us to detect active orthodoxal R-M systems in population, horizontal gene transfer between different species and viruses with wide specificity.

[1]   Ershova A. S. et al. (2105) Role of restriction-modification systems in prokaryotic evolution and ecology //Biochemistry (Moscow). – T. 80. – №. 10. – C. 1373-1386.
[2]   DeMaere M. Z. et al. (2013) High level of intergenera gene exchange shapes the evolution of   haloarchaea in an isolated Antarctic lake //Proceedings of the National Academy of   Sciences. – T. 110. – №. 42. – C. 16939-16944.
[3]   Tschitschko B. et al. (2015) Antarctic archaea–virus interactions: metaproteome-led analysis of   invasion, evasion and adaptation //The ISME journal.

[4]   Karlin S., Cardon L. R. (1994) Computational DNA sequence analysis //Annual Reviews in   Microbiology. – T. 48. – №. 1. – C. 619-654.