

## **Avoidance of recognition sites of restriction-modification systems as bacteriophage anti-restriction strategy**

Ivan S. Rusinov

*Lomonosov Moscow State University, Leninskiye gory 1, bldg. 40, Moscow 119992, Russia.*

*isrusinov@gmail.com*

Anna S. Ershova, Anna S. Karyagina

*Lomonosov Moscow State University, Leninskiye gory 1, bldg. 40, Moscow 119992, Russia;  
N.F. Gamaleya Center of Epidemiology and Microbiology, Gamaleya street 18, 123098 Moscow, Russia;  
Institute of Agricultural Biotechnology, Timiryazevskaya street 49, 127550 Moscow, Russia.*

*asershova@gmail.com, akaryagina@gmail.com*

Sergey A. Spirin, Andrei V. Alexeevski

*Lomonosov Moscow State University, Leninskiye gory 1, bldg. 40, Moscow 119992, Russia;  
Institute of System Studies, RAS, Nakhimovsky Prospekt 36, bldg. 1, 117218 Moscow, Russia.*

*sas@belozersky.msu.ru, aba@belozersky.msu.ru*

Restriction-Modification (R-M) systems are a kind of immune system of prokaryotes protecting them from invasion of foreign DNA, for example, DNA viruses. An R-M system is specific to a certain short DNA sequence, called recognition site. Prokaryotic viruses have been shown to avoid some short DNA sequences, known to be recognition sites of R-M systems, in their genomes. The avoidance is considered to be one (among many others) of anti-restriction strategies of bacteriophages [1]. However, any systematic study of restriction site avoidance in genomes of prokaryotic viruses based on the actual amount of data has not been published.

We analyzed restriction site avoidance in genomes of 2069 prokaryotic viruses. We used the relative abundance measure, suggested by Burge and co-authors [2], to detect a discrepancy between the observed frequency of a recognition site in a genome and its expected frequency under the assumption of absence of selection against the site in the genome. Genomes of eukaryotic viruses were used as negative control because they meet no R-M systems during their life-cycle (except some *Chlorella* viruses that encode R-M systems and were excluded

from the list). DNA bacteriophages demonstrate significant restriction site avoidance unlike RNA phages which are statistically indistinguishable from the negative control.

However, DNA bacteriophages commonly avoid only recognition sites of Type II R-M systems (excluding IIG subtype) while sites of other Types are avoided in scattered instances (see Fig. 1). It could indicate, for example, that bacteriophages have other widespread anti-restriction strategies targeting Type I, IIG, and III R-M systems.

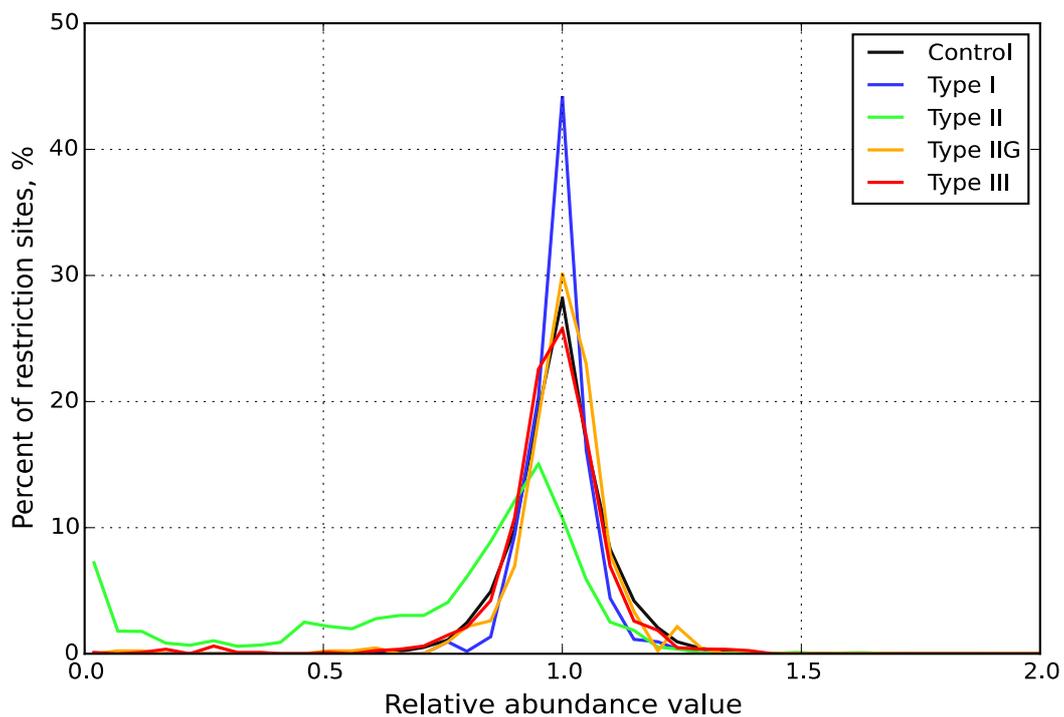


Fig. 1. Histograms of the relative abundance measure for recognition sites of R-M systems of different Types.

Temperate and lytic bacteriophages manifest different trends of Type II restriction site avoidance. Lytic phages more often (11.9% of sites) completely eliminate restriction site occurrences from their genomes than temperate phages (2.0% of sites). It might be caused by a long-term prophage stage during which a temperate bacteriophage shares host selective pressure affecting oligonucleotide composition of its genome. In support of the hypothesis, genomes of temperate bacteriophages statistically indistinguishable from equal (in length)

fragments of genomes of their hosts in terms of the used method of estimation of abundance of restriction sites.

We also demonstrated that coliphages of Myoviridae family encoding DNA-hydroxymethylase do not seem to develop restriction site avoidance (no significantly avoided sites was found), while the related phages without the gene avoid 73.7% of known host R-M system recognition sites. It allows us to suppose that at least some cases, when active R-M system is known in a bacterium but any avoidance of its recognition site was not detected in genomes of its bacteriophages, should be explained by other anti-restriction strategies of the phages.

1. M.R. Tock, D.T. Dryden (2005) The biology of restriction and anti-restriction. *Curr Opin Microbiol*, **8**: 466–472.
2. C. Burge, A.M. Campbell, S. Karlin (1992) Over- and under-representation of short oligonucleotides in DNA sequences. *Proc Natl Acad Sci USA*, **89**: 1358–1362.