High-throughput analysis of type I-E CRISPR adaptation in *E. coli*

E.E. Savitskaya

Skolkovo Institute of Science and Technology, Moscow region, Russia, Institutes of Molecular Genetics and Gene Biology of the Russian Academy of Sciences, Moscow, Russia, savitskaya.e@yandex.ru

E.A. Semenova

Waksman Institute for Microbiology and of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA, semenova@waksman.rutgers.edu

A.V.Strotskaya

Institute of Molecular Genetics and of the Russian Academy of Sciences, Moscow, Russia, Moscow, Russia, sandrynn5@gmail.com

K.A. Datsenko

Waksman Institute for Microbiology and of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA, Purdue University, West Lafayette, IN, datsenkk@purdue.edu

K.V. Severinov

Institute of Molecular Genetics and Gene Biology of the Russian Academy of Sciences, Moscow, Russia, Waksman Institute for Microbiology and of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA, Central Research Institute of Epidemiology, Russian Inspectorate for Protection of Consumer Right and Human Welfare, Moscow, Russia, severik@waksman.rutgers.edu

CRISPR is a prokaryotic genetic system that provides acquired immunity against plasmids and phages. This system is comprised of a cluster of short repeats (24-47bp long), interspersed by similarly sized non repetitive sequences (called spacers), and a set of CRISPR-associated (*cas*) genes. The CRISPR system is abundant among prokaryotes, and computational analyses show that CRISPRs are found in ~40% of bacterial and ~90% of archaeal genomes sequenced to date. CRISPR systems are divided into several subtypes, based on the composition of the *cas* genes and the repeat sequences.

It was shown that the CRISPR locus is transcribed into a single RNA transcript, which is then further cleaved by the Cas proteins to generate smaller CRISPR RNA units, each including one targeting spacer. These units, complexed with additional Cas proteins, then interfere with

the incoming foreign genetic material by complementary base-pairing with nucleic acids of the foreign element, leading to its cleavage and degradation. In response to phage (or plasmid) infection bacteria integrate new spacers identical to the phage genomic sequence (termed proto-spacer), resulting in CRISPR-mediated phage resistance. Both processes, CRISPRS interference and new spacers adaptation, are not understood completely in molecular terms.

Here, we analyzed *E. coli* CRISPR spacers acquired at conditions of overexpression of type I-E CRISPR/Cas system components and revealed by high-throughput sequencing. The uncovered determinants of new spacers choice and their impact to bacteria-phage coevolution will be discussed.