Autoimmune primed CRISPR adaptation in I-E and I-F systems: comparative analysis of new spacer selection mechanisms

Ekaterina Savitskaya[1,2], Ekaterina Semenova[3], Daria Vorontsova[1,4], Sofia Medvedeva[1], Maria Logacheva[5], Konstantin Severinov[1,2,3,4]

[1]*Skolkovo Institute of Science and Technology, Skolkovo Russia*

[2]*Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia*

[3]*Waksman Institute of Microbiology, Rutgers, the State University of New Jersey, Piscataway, NJ 08854,USA*

[4]*St-Petersburg State Polytechnical University, St. Petersburg, Russia*

[5]*Moscow State University, Moscow, Russia*

Prokaryotic adaptive immunity CRISPR-Cas systems consist of CRISPR (Clusters of Regularly Interspersed Short Palindromic Repeats) DNA arrays and *cas* (*C*RISPR *a*ssociated) genes. The CRISPR locus is composed of an array of identical repeats separated by spacers of common length but variable sequence. Spacers are acquired in a process termed CRISPR adaptation. The adaptation occurs when foreign DNA appears in the cell and leads to expansion of CRISPR array. The short motif recognized in target DNA called PAM marks those fragments (protospacers) in the foreign DNA that mainly will be incorporated in CRISPR array as spacers. The presence of spacers derived from foreign DNA allows CRISPR systems to protect cells from foreign DNA by specifically targeting it with ribonucleoprotein complexes containing Cas proteins and crRNAs transcribed from CRISPR array in a process called CRISPR interference. So called primed adaptation couples the CRISPR interference and CRISPR adaptation, as spacer acquisition is strongly facilitated in the presence of partial complementarily between a preexisting spacer and a protospacer. This process requires all *cas* proteins and leads to distinct strand bias, i.e., most of the of spacers acquired during primed adaptation originate from the strand where the primed protospacer is located. Another hallmark of primed adaptation is an increased bias towards the use of protospacers with PAM as a source of spacers.

Targeting the CRISPR-Cas system against protospacers in bacteria's own genome leads to double stranded breaks and arrest of cell division and is accompanied with

adaptation, allowing one to analyze the principles of new spacer choice in the course of primed adaptation. Two auto-targeting model systems were created: one for *E. coli* I-E CRISPR-Cas system and the second for *P. aeruginosa* I-F system, whose components were heterologously expressed in *E. coli* lacking its own CRISPR-Cas system. For both systems spacers acquired during primed adaptation were sequenced by Illumina and mapped to the auto-targeted genome. In both systems, new spacers were acquired from an extended area around the targeted protospacer with efficiency of protospacer choice dying out as the distance from targeted site increased. A clear strand bias was observed, which, however, was different upstream and downstream of the targeted site. For *E. coli* system, spacers coming from the region upstream of protospacer mapped to the non-targeted strand and spacers coming from the downstream region mapped to the targeted strand. Strikingly, the strand bias in *P. aeruginosa* system was reversed: spacers coming from the upstream region mapped to the targeted strand, while spacers coming from the downstream region mapped to the non-targeted strand. The whole genome high throughput sequencing of *E. coli* cells undergoing primed adaptation reveled a large (more than thousands of base pairs) gap in the targeted region whose size matched the area from which new spacers originated. Taking together our data imply a possible mechanism of primed adaptation where products of CRISPR interference serve as a source of material for new spacer acquisition.