

## Comparison of evolutionary rates of the regions and nucleotide substitutions in the *Allium* plastomes

D.O. Omelchenko<sup>1,2\*</sup>, A.A. Krinitsina<sup>1,2</sup>, M.D. Logacheva<sup>1,2</sup>, M.S. Belenikin<sup>3</sup>, E.A. Konorov<sup>1,4</sup>, S.V. Kuptsov<sup>1</sup>, A.P. Seregin<sup>1</sup>, A.S. Speranskaya<sup>1,2</sup>

<sup>1</sup> Lomonosov Moscow State University, Moscow, Russia

<sup>2</sup> Skolkovo Institute of Science and Technology, Moscow Region, Russia

<sup>3</sup> Moscow Institute of Physics and Technology, Moscow Region, Dolgoprudny, Russia

<sup>4</sup> Vavilov Institute of General Genetics, Moscow, Russia

\* e-mail:

**Key words:** *Allium*, evolution, high-throughput sequencing, chloroplast genome

**Motivation and Aim:** Genus *Allium* includes several economically important species, some of them of the oldest cultivated crops, along with some rare and endemic species. Still, complete sequences of *Allium* genomes are majorly underrepresented (only *A. cepa*, *A. sativum* and *A. obliquum* has complete chloroplast DNA records in GenBank). Modern molecular taxonomy of the genus is based on nuclear internal transcribed spacers (ITS) sequences analysis. Here we performed sequencing, assembly and comparative analysis on evolution and diversity of the complete sequences of the chloroplast genomes of *Allium* species, representing all three evolutionary lines of this genus.

**Methods and Algorithms:** The plastomes of 11 *Allium* species (*A. ursinum*, *A. paradoxum*, *A. zebdanense*, *A. victorialis*, *A. macleanii*, *A. fistulosum*, *A. nutans*, *A. platyspathum*, *A. obliquum*, *A. shoenoprasum*, *A. pskemense*) were sequenced on the Illumina MiSeq and assembled *de novo* (using various assemblers, including Velvet and MIRA4). Sequences were annotated by DOGMA and GeSeq, followed by manual review and correction of annotated features where necessary. Phylogenetic analysis was performed with Mr. Bayes and phangorn R package. Length of branches was calculated, and natural selection was detected by PAML and HyPhy. Evolutionary rates were calculated using Erable.

**Results:** Besides expected differences represented by small indels in intergenic spacers, there are conspicuous characteristic large deletions and nonsynonymous substitutions in some genes in the *Allium* plastomes, some of them specific to the evolutionary lines. *A. paradoxum* plastome showed most interesting results that differ it from other *Allium* species – 4.9 kbp inversion of *rpl32-ndhE* region with impairment of *ndhG* and *ndhF* genes in SSC. Evolutionary analysis showed that some genes commonly used for phylogenetic analysis (e. g., *matK*) have less constant evolutionary rate within the genus *Allium* compared to intergenic regions.

**Conclusion:** Analysis of *Allium* plastomes revealed prominent species specific and even evolutionary line specific differences in sequence and gene set. We assume the effect of evolution on various parts of *Allium* plastomes differs in character between species of this genus.

**Acknowledgements:** Supported by the RFBR grant No. 18-04-01203.