NGS data processing method for the mixture of chloroplast and mitochondrial DNA of barley

A. Makarevich*, O. Pankratov, M. Sinyavskaya, N. Lukhanina, A. Shymkevich,
A. Liaudansky, I. Goloenko, N. Danilenko, O. Davydenko
Institute of Genetics and Cytology, National Academy of Sciences of Belarus, Minsk, Belarus
* *e-mail: bio.makarevich@gmail.com*

Key words: NGS, data processing, chloroplast genome, mitochondrial genome, barley phylogeny

Motivation and Aim: Phylogenetic connections between cultivated barley and its wild forms are insufficiently explored field in evolution genetics [1]. Chloroplast and mitochondrial genome sequencing was performed for wild and domesticated barley lines. The fraction of chloroplasts obtained by differential centrifugation was used for DNA extraction. Such a DNA sample contains an admixture of mitochondrial DNA, which allows studying both chloroplast and mitochondrial genomes. However, the genome assembly is complicated by presence of extended areas of homology within chloroplast and mitochondrial genomes as well as between them. Thus, the aim of this research was to develop the NGS-data processing method, which allows assembling complete sequences of each genome from the mixture of chloroplast and mitochondrial DNA.

Methods and Algorithms: The algorithm of sequencing data processing included the following steps: trimming of row reads (Trimmomatic); aligning reads to the "double" reference, containing full sequences of chloroplast and mitochondrial barley genomes (Bowtie2); obtaining mapping statistics (bash scripts, BCFtools); alignment visualization (Tablet); generating VCF files (Samtools); filtering VCF files (VCFlib). The algorithm was tested on artificial Illumina reads synthesized using the ART program. Ultimately, VCF files containing all polymorphic loci of the chloroplast and mitochondrial genomes were obtained.

Results: Testing of the developed algorithm showed its applicability for the mixtures of barley chloroplast and mitochondrial DNA with different ratios of concentrations.

Conclusion: The developed algorithm allows obtaining complete sequences of chloroplast and mitochondrial barley genomes from the mixtures of plastid and mitochondrial DNA using the next-generation sequencing data. Modifications of this method can be used to study the organelle genomes of other cereals.

Acknowledgements: The study was conducted within the framework of the State Program of Scientific Research "Biotechnologies 2016–2020".

References
1. Pankin A., von Korff M. (2017) Co-evolution of methods and thoughts in cereal domestication studies: a tale of barley (*Hordeum vulgare*). Current Opinion Plant Biology. 36:15-21.