# Sequencing, de novo transcriptome assembly and differential expression analysis of flax genotrophs

Maxim Belenikin

*Research Institute of Physico-Chemical Medicine, 119435 Moscow, Russia,* camd@niifhm.ru

N. Melnikova

*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 119991 Moscow, Russia*

A. Speranskaya, A. Krinitsina

*Moscow State University, Faculty of Biology, 119991 Moscow, Russia*

N. Bolsheva, V. Lakunina, O. Rachinskaya, A. Dmitriev, M. Darii, A. Kudryavtseva, A. Zelenin, O. Muravenko

*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 119991 Moscow, Russia*

Flax (*Linum usitatissimum* L.) is important agricultural crop. Flax stem fiber and seed oil have multiple industrial applications. Flax undergoes heritable changes in phenotype and genotype in response to the growth environment [1]. Flax lines in which stable inherited changes were observed were named genotrophs. Very little is known about mechanisms of genetic changes during flax development under defined environment conditions. The comparison of mRNA expression profiles was performed for flax cultivar Stormont Cirrus grown under normal (1), inadequate (2) and excessive (3) nutrition using the RNA sequencing with genome analyzer Illumina GAIIx. Library preparation and sequencing processes for three states we performed in according to standard Illumina`s protocols. A total of 13.6M (1), 10.6M (2), and 18.6M (3) raw 150 bp reads were obtained from high-throughput sequencing of flax transcriptome. The transcriptome assembly was performed using Trinity de novo transcriptome assembler [2]. The total length of transcripts was 139.74Mbp, max transcript length 10511bp, the total number of transcripts/length cutoff/mean transcript length were: 100395/200bp/1391bp; 67837/500bp/1913bp; 51993/1000bp/2273bp. Quantification, and identification of differentially expressed transcripts were performed by RSEM [3] and edgeR [4]. About 3100 transcripts with RPKM equal or more fifty were separated for next tblastx analysis; 90.4% transcripts were annotated. A number of the most interesting transcripts were selected. Thus, during current study for flax genotroph cultivar Stormont Cirrus we performed: (I) cDNA sequencing at

three nutrition conditions using Illumina GAIIx platform; (II) de-novo transcriptome assembly by Trinity software; (III) differential expression analysis. It gives new information about the processes occurring while the plants were growing under different conditions and the role of the environment in generating adaptive mutation.

The work was financially supported by grant of the President of Russian Federation MK-6205.2012.4, RFBR research project No. 12-04-01469a and RFBR research project No. 13-04-01770.

1. Y.Chen et al. (2009) An environmentally induced adaptive insertion event in flax. *Int. J. Genet. Mol. Biol.*, **1**(3):38-47.

2. M.G.Grabherr et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome, *Nature Biotechnology*, **29**:644-652.

3. B.Li, C.N.Dewey (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, *BMC Bioinformatics*, **12**:323.

4. M.Robinson, A.Oshlack (2010) A scaling normalization method for differential expression analysis of RNA-seq data, *Genome Biology*, **11**:R25.