

iTraVA: transcriptome map of *Arabidopsis thaliana* alternative isoforms

A.V. Klepikova^{1*}, A.M. Kasianova^{1,2}, T.A. Isaeva, M.D. Logacheva^{1,2}, A.A. Penin¹

¹ Vavilov Institute of General Genetics, RAS, Moscow, Russia

² Center of Molecular and Cellular Biology, Skolkovo Institute of Science and Technology, Moscow, Russia

* annklepikova@gmail.com

Motivation and aim: Splicing is a key event in mRNA maturation. Alternative isoforms play crucial role in the development and environmental interactions of both animals and plants. Alternative isoforms differ between organs, tissues and developmental stages, guiding their functional differences. The understanding of splicing complexity in plants is far from complete. Thus, we aimed to create a comprehensive isoform-based transcriptome map across a variety of *Arabidopsis thaliana* tissues and present it in a new public database iTraVA: isoform Transcriptome Variation Analysis.

Materials and methods: 29 samples of *A. thaliana* organs and tissues in normal conditions and under various stress treatment were collected and subjected to cDNA sequencing on Oxford Nanopore Technologies platform. Isoforms expressed in each sample were reconstructed using FLIC tool developed in the laboratory. Isoform properties were extensively studied.

Results: We created an isoform-based transcriptome map covering 29 samples of *A. thaliana*. In total, more than 158 thousand isoforms were detected in the whole transcriptome map belonging to 20 thousand genes. The majority of isoforms had predicted ORF longer than 30 a.a. We found that the median number of isoforms per gene was seven considering the whole amount of data, and varied between two and four in individual samples. As FLIC detects transcription start sites (TSSs) and polyadenylation sites (PAs) as regions we analyzed the distributions of TSS and PA length and found them to be bimodal. We explored alternative splicing events both across the whole map and in individual samples and analyzed the difference between genes with certain types of events as well as the inequality of TSS and PA usage. We found that in many cases isoforms had more narrow expression patterns than their corresponding gene. The structure and expression levels of reconstructed isoforms were presented in the newly developed database iTraVA.

Conclusions: We created an isoform-based transcriptome map of *A. thaliana* and explored isoform properties in greater detail. Our transcriptome map is available in the database iTraVA.

Funding: The study was supported by Vavilov Institute of General Genetics, project No. FFRW-2024-0003.