**Bioinformatics tools for whole transcriptome sequencing data analysis in leukemia patients with complex structural variants**

Jakub Hynst1,2, Karla Plevova1,2,3, Lenka Radova1,Vojtech Bystry1, Karol Pal1,2, and Sarka Pospisilova1,3

*1 Central European Institute of Technology, Masaryk University, Brno, Czech Republic*

*2 Faculty of Medicine, Masaryk University, Brno, Czech Republic*

*3 University Hospital Brno, Brno, Czech Republic*

**Background.** Recently, extensive genome rearrangements known as chromothripsis have been identified in several cancer types. They lead to complex structural variants (cSVs) causing changes in gene expression profile and formation of *de novo* fusion genes. At RNA level, functional impact of cSVs can be studied using whole transcriptome sequencing (total RNA-Seq). Nevertheless, bioinformatics analysis of RNA-Seq data, especially in cases with cSVs, still remains challenging and development of bioinformatics tools is necessary.

**Methods and results.** We designed bioinformatics tools for RNA-Seq data analysis consisting of two parts: The first pipeline enables analysis of differential gene expression in a biologically heterogeneous sample set. The other pipeline serves for *de novo* fusion gene identification with a special attention to false positive call filtering based on combination of several fusion gene callers. We applied both tools on RNA-Seq data obtained from a cohort of ten patients with chronic lymphocytic leukemia (CLL). Results of differential gene expression were concordant with a parallel transcriptomic array experiment, which confirmed analytic abilities of our method. We also showed that fusion gene detection approach was capable to identify true positives efficiently, as parallel experiment on genomic arrays provided concordant results.

**Conclusion.** Applying our tools on real RNA-Seq dataset from CLL patients, we proved that our approach generates consistent results with other genomics analytical techniques. The tools provided clues for studying biological consequences of cSVs with implications for clinical outcome and management of cancer patients. The tools are applicable for addressing also other research questions in different contexts.

Supported by AZV-MZCR 15-31834A, CEITEC2020 LQ1601, MZCR-RVO 65269705, and MU Medical Faculty MUNI/A/0968/2017.