Reviewer assessment of the doctoral thesis „Transcriptomic Characterization Using RNA-Seq Data Analysis“

Author: Ing. Layal Abo Khayal, Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology

Doctoral study program: Electrical Engineering and Communication

Field of study: Biomedical Electronics and Biocybernetics

Reviewer: prof. PharmDr. Petr Babula, Ph.D., Department of Physiology, Faculty of Medicine, Masaryk University, Kamenice 5, CZ-625 00 Brno

The doctoral thesis aims to develop and innovate approaches in the analysis of large volumes of data that are usually captured in RNA-Seq experiments. These new and innovative approaches are usually of enormous demands, especially regarding the need to use integrated workflows. Transcript identification and the quantification of gene expression have been distinct core activities in molecular biology ever since the discovery of RNA's role as the key intermediate between the genome and the proteome. In view of the biological information, it is necessary to determine not only the level of gene expression, but also possible alternative splicing and determination of long non-coding RNA, because these processes finally lead to the final product of gene expression - protein.

In the theoretical introduction, the author introduces the methods of RNA isolation, control of RNA quality, preparation of cDNA library based on RNAs in a sample, different RNA-Seq platforms and RNA-Seq Applications. The goals of the doctoral thesis are clearly defined and described. For the work, the author used RNA-Seq data from differentiated osteoblasts. Primary calvaria osteoblasts were harvested according to the published protocol, seeded, and harvested again, but at different differentiation points (days 0, 3, 6, and 12). After it, RNA was isolated using chloroform/phenol extraction and it integrity was verified. Subsequently, transcriptomic data were analysed. The analyses are presented in the chapter RNA-Seq Data Analysis Workflow and include differential gene expression analysis, gene ontology analysis enrichment, usage of differential exons, and identification of long non-coding RNA. The work is clearly organized into individual chapters, which are well organized in a logical way. The author works with a total of 141 references, which are mostly recent. The topic of the dissertation is very interesting, inspiring, scientifically up to date, and in some ways innovative, when it reaches a wide range of fields, from bioinformatics and biology to molecular biology. The author designed and implemented her own approaches to RNA-Seq data analysis. She uses the computational pipeline adequate for the analysed RNA-Seq data.

To each of the obligatory points:

1. Does the doctoral thesis correspond to the subject of the dissertation and does it correspond to the current state of science?

The subject of the thesis corresponds to the field of dissertation and is very current in terms of the up to date state of science. In addition, the subject of the thesis assumes that it will be possible to further develop the work.

2. Does the work show original beneficial part?
The work shows the original beneficial parts, as I mentioned above.

3. Has the core of the dissertation been published at the necessary level?
Partial results passed through the peer review process and were published.

4. Does the applicant's scientific activity indicate that he is a scientist worker?
The applicant's list of scientific activities shows that she is a scientist. This is evidenced by the fact that she has published (or which are under review process) 4 contributions, of which 3 times in the position of the first author.

Reviewer has following specific comments and questions:

1. Why did the authors use phenol/chloroform extraction of RNA? This method has many disadvantages. In addition, there are many other, far more suitable methods for the isolation of RNA. Can the author introduce the methods, which are friendlier to RNA itself during its isolation?
2. Why did the authors choose osteoblasts for the analysis? Is there any special interest in these cells?
3. The manuscript "Impaired proteoglycan glycosylation...." – where is contribution of the author of the doctoral thesis. Can she specify this?
4. RNA-Seq has a wide variety of applications, but no single analysis pipeline can be used in all cases. The major analysis steps include pre-analysis, core analysis, and advanced analysis. Pre-analysis includes experimental design, sequencing design, quality control steps, core analyses include transcriptome profiling, differential gene expression, and functional profiling, and advanced analysis includes visualization, other RNA-Seq technologies, and data integration. Can the author introduce (briefly) own contribution in these areas?
5. Can the author introduce possibilities of the integration of RNA-Seq data with other types of genome-wide data (connecting the regulation of gene expression with specific aspects of molecular physiology and functional genomics – combination with DNA-Seq, Pairwise DNA-methylation and RNA-seq integration, the combination of RNA-Seq and transcription factor chromatin immunoprecipitation sequencing (ChIP-Seq), integration of RNA-Seq and miRNA-Seq data, integration of RNA-Seq with proteomics).
6. What is the opinion of the author on modern approaches in RNA-Seq (single-cell RNA-Seq, long-read sequencing) and possible processing of data from these analyses?

Finally, I note that the doctoral thesis fully meets the requirements of this type of work in the given field. It points to the ability of the author to work in a scientific team, to use highly sophisticated methods, to process, interpret and present corresponding results. In accordance with the relevant paragraphs of act No. 137/2016 Coll. (Act on Higher Education No. 111/1998 Coll.) I recommend acceptance of the doctoral thesis for the defence and the title of Ph.D. after its successful defence.

prof. PharmDr. Petr Babula, Ph.D.
Department of Physiology, Faculty of Medicine, Masaryk University

Brno, March 12, 2018