



FACULTY OF ELECTRICAL department
ENGINEERING of biomedical engineering
AND COMMUNICATION

prof. Ing. Ivo Provazník, Ph.D.
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology
Technická 12
61200 Brno
Czech Republic

Review of habilitation thesis

Evolution of Genes and Genomes: Algorithms and Models

written by Mgr. Tomáš Vinař, PhD.

Tomáš Vinař, PhD. has submitted a habilitation thesis entitled "Evolution of Genes and Genomes: Algorithms and Models". There have been attached the following documents to the thesis:

- Personal scientific contribution assessment
- Professional Curriculum Vitae
- Overview of didactic activities
- List of the most important publications and citations
- List of other scientific and professional publications
- List of other citations
- Originality check report

The thesis is focused on bioinformatics (computational biology) namely on recent research and development of new algorithms, data structures and models for analysis of biological sequences and for comparative genomics. The following problems are raised: application of hidden Markov models to gene search, problems of inferences of the most probable annotations of de novo sequenced genomes, searching for orthologous genes in unannotated genomes, searching for orders of genes in sequences, and others. The described algorithms are based on statistical and probabilistic methods using various types of models. The theme of the habilitation thesis correspond to the field of habilitation.

The thesis contains 172 pages of text written in English including a list of 60 relevant scientific publications and is divided into five chapters and seven appendices. The first part of the thesis (chapters 1 to 5, 40 pages in total) represents introduction to the mentioned field. It also introduces a reader to particular topics and summarizes all relevant results. The thesis is completed by appendices with full texts of 7 research papers of the author published in major scientific journals and proceedings of international conferences.

The thesis covers three major areas:

1. Evolution of protein coding genes
2. History of duplicated gene clusters
3. Genome rearrangements

Chapters 2, 3 and 4 provide short but clear introduction to these three complex problems and refer to the relevant papers published by the author. The papers themselves represent detailed studies with important and useful outcomes and provide a number of possible scenarios for future work.

The main topic of the thesis is not new in the field of bioinformatics. The selected approaches have also been on scene for some years. However, I find orientation of the thesis highly up-to-date and important. The author correctly mention the constantly changing environment of sequencing technology when significantly different devices are used in each new generation of sequencers. From this point of view, I also find the topic of the thesis recent, perspective, and internationally competitive. This is also proved by selection of publications in Bibliography chapter that contains (among others) references to relevant international journal papers from period of 1969 – 2014. Further, the algorithms authored or co-authored by Dr. Vinař and presented in the thesis contain unique approaches and methods that undoubtedly move the field ahead and contribute importantly to recent bioinformatics.

Question for discussion

Recent sequencing technologies tend to provide large-size whole genome data at low cost and high speed. This will lead to fast production of large data sets and push bioinformatics into big-data field. At the same time, approaches to sequence assembly change as well in term of level of data fragmentation. Could you explain how your methods, algorithms, and data structures should be re-designed to processing data from Oxford nanopore technology as an example of near-future technology producing up to 200 Kb long reads in a real-time stream? This might be a widely-used approach that needs different data processing.

Conclusion

Based on analysis of the habilitation thesis entitled "Evolution of Genes and Genomes: Algorithms and Models", its educational, professional and scientific impact, originality of the presented results, scientific merit and impact of the author research, I can state that the thesis **meets** all general requirements for the habilitation works in the field of bioinformatics. Thus, I **recommend** Mgr. Tomáš Vinař, PhD. can obtain the scientific-pedagogical title

"docent – associate professor"

in the branch of Informatics (9.2.1).

Brno, 25.12. 2016