UNIVERSITÄT LEIPZIG

Fakultät für Mathematik und Informatik
Institut für Informatik
Professur für Bioinformatik
Prof. Dr. Peter F. Stadler

Universität Leipzig, Bioinformatik, Härtelstraße 16 – 18, 04107 Leipzig

Bearbeiter:
Sven Findeiß
Telefon 0341 97-16690
Telefax 0341 97-16679
findeiss@bioinf.uni-leipzig.de

28. November 2019

Official Review
of the thesis entitled "Development of bioinformatic approaches for studying the molecular mechanisms of the chronic human diseases" by Arsen Arakelyan for the Doctor of Sciences degree in 00.00.03 Molecular and Cellular

Universität Leipzig
Fakultät für Mathematik und Informatik
Institut für Informatik
Professur für Bioinformatik
Härtelstraße 16 – 18
04107 Leipzig

Telefon
+49 341 97-16690

Fax
+49 341 97-16679

E-Mail
peter.stadler@bioinf.uni-leipzig.de

Web
www.bioinf.uni-leipzig.de

Postfach intern
505001

Dr. Arsen Arakelyan submitted a thesis with the title "Development of Bioinformatic Approaches for Studying the Molecular Mechanisms of Chronic Human Diseases". It comprises 5 single-author papers in international journals, 15 reviewed manuscripts arising from national and international collaborations as well as 3 book chapters. New options for prognosis, diagnosis, and treatment of non-communicable diseases became available with the advent of new high-throughput omics-technologies. In his thesis, Dr. Arakelyan developed and applied a series of bioinformatics methods and tools to process the new type of high-dimensional data and to demonstrate their impact in practical data analysis in the context of different diseases. The development of prognostic, diagnostic and therapeutic methods for chronic noncommunicable diseases has become an important component of molecular medicine with a rapidly increasing impact on healthcare. Computational approaches, in particular, the understanding of regulatory mechanisms involved in the origins and development of this broad class of diseases is a rapidly growing topic in computational biology and bioinformatics, to which Arsen Arakelyan has made significant contributions in four areas:

1) He developed combinatoric analysis schemes for gene expression analysis, particularly for classifying genes according to their impact in case-versus-control comparisons.

2) Dr. Arakelyan essentially contributed to the development of pathway-based analysis schemes which take into account the topology and direction of gene-gene interaction. His pathway signal flow (PSF) is an essential contribution to the field. KEGGParser and CyKEGGParser are tools to parse, visualize and analyze many aspects of the metabolic and regulatory databases. The latter provides access to the methods in Cytoscape and is widely used in the community with nearly 30000 downloads. In subsequent work, the techniques were used to show,

among other results, signaling pathways largely are robust to mutations that alter protein-protein interactions due to complex branched topologies. On the other hand, pathways may contain individual hub or bottleneck genes that are highly susceptible to perturbations.

3) A. Arakelyan applied his PSF method in large scale meta-analyses on lung and autoimmune diseases by developing a machine learning approach which enabled him to sort pathway contributions and their changes in the diseases under study. In this work, Arsen Arakelyan also features a series of detailed work focussed on specific disease complexes. His systems-levels analysis of interstitial lung diseases, for instance, identified three new subtypes distinguished by differences in the links to immune/inflammatory response and fibrotic tissue remodeling. This is distinct from the involvement of cell proliferation and metabolism in lung cancers. Autoimmune and autoinflammatory diseases are distinguished by similar differences in the involvement of immune and inflammatory response-related pathways. In this line of work, Arsen Arakelyan also established that autoinflammatory processes play an important role in several autoimmune diseases. An interesting result is that different causal mutations in monogenic autoinflammatory syndromes lead to a highly divergent gene expression profile, while at the level of pathways one observes a coherent picture of deregulation.

4) Finally, a quite different field application is post-traumatic stress disorder (PTSD). Here, Arsen Arakelyan identified informative prognostic markers in gene expression profiles and peripheral blood, which -- in late stages of the disease -- correlate with the main neuropsychiatric parameters.

Meanwhile, I would like to point a few minor issues worth to be addressed. Dr. Arakelyan states that the tools developed could be used with other -omics data rather than transcriptome; however, there are no examples for that in the thesis. Another question if PSF accounts for stochiometric information of protein-protein interactions within a pathway. Finally, in the analysis of mutation effects on pathway activation, only disruptive mutation cases were considered. However, it is known that in many cases mutations cause a range of changes in the binding interaction strength and it is interesting how this information could be integrated into PSF analysis.

In summary, Arsen Arakelyan's research contains several important contributions to computational and molecule medicine. Both the computational tools and the detailed case studies have a lasting impact on their field. Correspondingly, much of his work is published in important international journals, among them leading publications in their fields such as Bioinformatics, Genome Medicine, or Journal of Pathology.

I therefore strongly believe that the thesis fully complies with the requirements of the Higher Certification Committee and I recommend awarding Arsen Arakelyan with Doctor of Sciences degree in "Molecular and Cellular Biology".

Kind regards,

Prof. Peter F. Stadler