Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Mikhail Moldovan

PhD Program: Life Sciences

Title of Thesis: Heritable modifications of transmitted biological information as possible sources of adaptation

Supervisor: Professor Mikhail Gelfand

Name of the Reviewer: Prof. Dr. Olga V. Kalinina

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	 **Date: 04-08-2022**

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

The dissertation of Mr. Moldovan reports his contributions in studying heritable modification that happen beyond the DNA level: adenine-to-inosine (A-to-I) RNA modification and post-translational protein modifications, in particular phosphorylation. He discusses the role of these modification in various evolutionary processes and shows that they could be in complex interplay with forces that act on fixation of DNA mutations. The abstract of the thesis presents the three major reported results: (1) positive selection on A-to-I edited adenines in RNA in soft-bodies cephalopods (coleoids); (2) epistasis-related clustering of such sites; and (3) clustering and positive selection on phosphorylated amino acid residues in protein sequences.

The dissertation is structured in six with appendices. Chapter 1 sets the field and the place of the presented work in it and lists the author's individual contributions. Chapter 2 gives the biological background and introduces the relevant terminology: *primary sites*, i.e. the sites of mutations; *secondary sites*, i.e. the sequence context of mutations; and *tertiary sites*, i.e. the molecular entities that are responsible for introduction of mutations. Then Mr. Moldovan spends considerable space on describing diverse evolutionary mechanisms that may be involved in the heritable information modifications on the RNA and protein levels and discussing the potential implications and possible analyses to verify them. I find this a very thorough and useful analysis, but I think a little more care could have been spent on proper definition of all terms (e.g., the word 'decodings' appears here for the first time, does not seem to be standard, and is never precisely defined).

Chapters 3 to 5 report three published studies where Mr. Moldovan is the first author.

In Chapter 3, a study on adaptive evolution of A-to-I mRNA editing in coleoids is reported. The study is performed by analyzing transcriptomes of four related coleoid species. The main results here are: (1) the editing level (percentage of edited transcripts) depends on the immediate sequence context, as well as on the local RNA structure; and (2) there appears to be positive selection for A-to-G substitutions at the edited sites. These are very important observations that contribute to our understanding of evolutionary processes. Again, my only criticism here is to an incomplete definition of notation: one has to guess what 'E' means in the context of 'E-to-G and/or G-to-E transitions' – I imagine that means the edited adenines, but that is only my assumption. A theoretical framework is provided at the end of the chapter, suggesting plastic adaptation or preadaptation as possible mechanisms, which bears an added value in the general context of the thesis.

Chapter 4 reports a study on clusters of A-to-I edited sites on coleoids using the same transcriptomics data. The main results here are: (1) consistent with previous observations, the editing sites cluster significantly along the RNA sequence forming longer chains of up to eight consecutive adenines, and editing levels in the adjacent sites are correlated; (2) there is an epistatic interaction between adenine editing probabilities at adjacent sites, which may be caused by common RNA structure; and RNA structure may bring distant editing sites closer together; and (3) editing is directional: downstream adenines are more probable to be edited, and this agrees with the probability of A-to-G mutations as evident from phylogenetic analysis and ancestral sequence reconstruction. In this chapter, I have to point out a redundancy I spotted in the Methods section: in my opinion, sections 4.2.2 and 4.2.9 can be placed together and merged. Also, I think Fig. 12B should be Fig 6B. Nonetheless, the presented results are of a major importance for understanding evolutionary processes involved.

Chapter 5 reports a study on a different subject: distribution of phosphorylation sites in proteomes of mammals. Here the main results are as follows: (1) phosphorylation label tends to be conserved in well-studied proteomes (human, mouse, rat), whereas the identity of the phosphorylated residue may differ; (2) phosphorylation sites tend to cluster with each other; and (3) there are distinct mutation patterns of amino acids in phosphorylation sites. The author here claims that for phosphorylated serines the

dominant mutation mode is to negatively charged amino acids, which would be in concert with the negative charge of the phosphate group that is attached during phosphorylation. However, from Fig. 13B I can see that mutations to lysine (a positively charged amino acid) are almost as prominent. I wish Mr. Moldovan could comment on this. (Also, the color coding in Fig. 13 is not explained in the figure legend.) Another result from this study is that mutation patterns are correlated with tissue-specific phosphorylation with sites that are phosphorylated in more tissues being more prone to mutations to negatively charged amino acids.

Finally, Chapter 6 conclude the dissertation and gives a brief outlook.

The methods presented in this dissertation clearly advance the research in the corresponding domain. They provide novel insights into laws of evolution. The reported results have been published in three first-authored peer-reviewed papers which is a remarkable output for an early career researcher. The above-mentioned criticism do not diminish the value of the presented research and the quality of dissertation, and can be easily addressed by the candidate.

The dissertation demonstrates a high degree of scientific and professional maturity of Mr. Moldovan. It witnesses that he has acquired versatile skills, which makes him capable of contributing in various domains both inside and outside of the academia.

I accept this thesis and recommend that the candidate defends it by means of a formal thesis defense.

Provisional Recommendation

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*