Thesis Changes Log

Name of Candidate: Evgeniia Alekseeva

PhD Program: Life Sciences

Title of Thesis: Evolutionary analysis of intrahost interaction between pathogens and adaptive immunity

Supervisor: Prof. Georgii Bazykin

The thesis document includes the following changes in answer to the external review process.

I am very grateful to the reviewers for their detailed study of my work and provided suggestions. Below are my point-to-point replies and corresponding changes of the Thesis text.

Doctor Alexander Bagaev

- *While Chapter 2, the literature review, is well-written and adequately referenced, it would be beneficial to expand the section discussing B cell clonal evolution analysis, particularly in the context of lymphomas.*

  Certainly, the application of evolutionary analysis to malignant B cell clonal lineages has the potential to provide valuable insights into the study of B cell lymphomas. I added a brief overview of such studies in the corresponding section of the literature review (*page 35*).

- *The results from the two chapters of the manuscript could be more interconnected, enhancing the overall coherence of the thesis.*

Thank you for this comment, it is complementary to comments of Prof. Khrameeva. Summing all these comments together I restructured several sections to make the thesis more coherent. First I replaced introductions devoted to result-related chapters into shortened paragraphs, which integrate described studies into the Thesis topic. I also joined two separate lists of conclusions into a single one, which conclude the storyline of the study of intrahost coevolution between pathogens and adaptive immunity.

- *It would be logical to include the B cell repertoire analysis in the second chapter as well.*

I agree that the first part of the study with general B cell repertoire analysis complements the ongoing part with the evolutionary analysis. However it comprises a chapter of another Skotech PhD Thesis, defended recently by my co-first coauthor Artem Mikelov. Thus, I did not include his part as he did not include mine to avoid confusion about repeating results.

- *The manuscript contains minor grammatical and proofreading errors that should be corrected.*

Unfortunately it definitely does. I tried to catch them again as much as possible.

Prof. Dmitry Ivankov

Thank you very much for the review, as far as I understand no suggestions to modify the Thesis text is provided.

Prof. Ekaterina Khrameeva

- *The literature review is followed by Chapter 3, which contains a slightly modified text of the paper (Mikelov et al. 2022). It looks a bit unusual to me that this chapter contains its own Introduction section. This Introduction section was apparently copy-pasted into the thesis from (Mikelov et al. 2022) with little modifications, which is ok, but it seems to be excessive because the thesis already has an Introduction section at the beginning and a detailed literature review. This additional Introduction section repeats what has already been described above. The same comment goes to the Introduction section at the beginning of the Chapter 4. I would recommend replacing it with some text linking Chapters 3 and 4.*

I agree that introductions of chapters 3 and 4 largely repeat a literature review. I replaced them with shortened paragraphs, helping to integrate these studies in the general storyline of the Thesis (*pages 48 and 70*) .

- *In Fig. 3.2G, it is not clear how the two-sided Mann-Whitney test was applied to calculate the statistical significance in this case. Fig. 3.2G presents two fractions while the Mann-Whitney test is applied to some distribution(s) of values. The procedure should be better described.*

Thank you very much for noticing this mistake in the figure description. Mann-Whitney test was applied just for figures B and F on this Fig. 3.2 panel. Statistical significance for Fig. 3.2G was calculated by Two proportions Z-test and was missed in the figure description. Now it is fixed (*page 58*).

- *In Fig. 4.1B, it is unclear what the legend Coverage Depth describes. It seems unrelated to the Fig. 4.1B. Probably, it is related to the Fig. 4.1D and should be moved there.*

I agree that the legend position is confusing on Fig. 4.1. Now it is moved closer to the D panel (*page 79*).

- *Is it possible to calculate the statistical significance in Fig. 4.3B,D?*

Figure 4.3 is obtained by my coauthors. Each bar represents the fraction of cells, responding on stimulation from the general number of cells used in the experiment. Since stimulations by ancestral and mutated peptides were independent and were done on separate samples with different numbers of overall T cells it is possible to calculate two-proportions Z test between them. To do so I requested corresponding numbers of responded cells and their overall counts and added significance on the figure with the description of the test in the figure's legend (*page 84*).

- *In the Conclusions chapter, the link between the two parts of the work should be better articulated. Ideally, this chapter should contain one list of conclusions (with clear links between them) instead of two separate lists. And I would recommend adding a small paragraph to the very end of this chapter, which would sum up both parts of the work. I.eA more general concluding paragraph is missing here.*

Indeed conclusions were not coherent enough. I restructured the whole section and joined two lists of conclusions in a single and more general one, reflecting the topic of the thesis (page 89-90).

Doctor Grigory Efimov

- *The main criticism relates to the central part of the study (Chapter 4). The limitation of the study is that the authors did not investigate the response to other SARS-CoV-2 epitopes that are known to be present in the patient's HLA class I alleles and thus did not show what fraction of SARS-CoV-2 specific CD8+ cells was affected by the accumulation of the studied mutations.*
  Another limitation of the paper is that the authors limit their pipeline by studying mutations that affect

predicted HLA binding, whereas it is known that there are SARS-CoV-2 mutations in CD8+ epitopes that do not affect epitope presentation but prevent epitope recognition by T cells (see Dolton et al. Cell 2023 for an example).

I definitely agree that our study has these and some other limitations. The estimation of the fraction of CD8+ T cells, from which viral population escapes, is rough and ideally requires an experiment with the mix of all possible changed epitopes in ancestral and derived states. Thus the level of immune escape in our study can be underestimated since we did not check epitopes with no prior information about their immunogenicity, even if they lost their binding due to predictions of NetMHCPan. Also we did not check epitopes, which did not lose their binding. Nevertheless altered amino-acid sequences still may prevent antigen recognition by corresponding TCR and result in immune escape. For the most part this is due to the fact that we were extremely limited in the amount of blood collected from the patient in an early (summer) time point for experiments with T cell stimulation. Those T cell clones, which responded to ancestral states of escaped epitopes in summer, were gone already in the later time point in winter. This can be both due to the shrinkage of corresponding T cell clones without antigen stimulation or due to hematopoietic stem cell transplantation, which should restructure the whole patient's TCR repertoire. Thus, at the time when this analysis was carried out, unfortunately it was already useless to collect blood material for such experiments. All these limitations are described in the Discussion section of the Chapter 4 (*pages 87-88*). Thank you for the link on a very relevant and interesting study, I added the link in the Discussion as well.

- *Minor comment: some abbreviations are missing from the list (MRCA, HSCT, PHBR, BR).*

Thank you for finding that out, abbreviations are added to the list now (*page 12-13*).

Prof. Mikhail Gelfand

- *It might be a good idea to proofread the text prior to the final submission, as it contains a considerable number of misprints, grammar and spelling errors, and incorrect choice of words e.g. in the Abstract: "immune repertoires sequencing technologies", "intrahost evolution of pathogens or pathological processes such as viruses or tumor cell lines are more frequently studied", "all these together", or, in the Acknowledgements, "approved my computational conclusions by lab experiments", or, in the Introduction, "the first part of the thesis (Chapter 2)" (should be "Chapter 3"), or (page 25) "cells, who does not get survival signal" and "sum diversity", or (page 30) "criterias".*

Thank you for finding that out, corresponding changes are introduced in the text (*pages 3, 7, 15, 16, 22, 29*).

- *More information about conferences should be provided (location, dates, form of presentation); "Moscow Conference of Computational Biology" in fact "International Moscow Conference on Computational Molecular Biology".*

More information about presentations on conferences is provided (*page 6*).

Prof. Richard Neher

Thank you very much for the review, as far as I understand no suggestions to modify the Thesis text is provided.