Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Artem Mikelov

PhD Program: Life Sciences

Title of Thesis: Dynamics of immunoglobulin repertoires in memory and antibody-secreting B cell subsets in health and disease

Supervisor: Associate Professor Dmitriy Chudakov

Name of the Reviewer: Prof. Dr. Katharina Röltgen

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: 24-03-2023**

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

Swiss TPH
Swiss Tropical and Public Health Institute

Prof. Dr. Katharina Röltgen
Head, Viral Immunology
katharina.roeltgen@swisstph.ch
Direct +41 61 284 86 40

PhD thesis evaluation report

It was a great pleasure to read Artem Mikelov's Doctoral Thesis: "Dynamics of immunoglobulin repertoires in memory and antibody-secreting B cell subsets in health and disease". Artem's work was focused on a detailed analysis of immunoglobulin heavy (IGH) chain repertoires in longitudinal peripheral blood samples collected from healthy subjects over the time course of one year. The goal of his research was to characterize specific features and dynamics of antigen-experienced B cell subsets in the absence of acute or chronic disease states to complement and expand a growing body of literature on these differentiated B cell subsets in response to pathologic conditions and vaccination. During his PhD thesis, Artem and his colleagues successfully tackled a major difficulty in the downstream analysis of IGH repertoires – the genotyping and inference of novel allelic Variable (V) and Joining (J) gene variants – by developing a new computational method addressing this challenge, called the MiStrainer algorithm.

1. *Brief evaluation of the thesis quality and overall structure of the dissertation:*

The PhD thesis has yielded high quality data on longitudinal IGH repertoires of antigen-experienced B cell subsets in healthy individuals. Computational analyses and preparation of figures illustrating the results were carried out in a highly professional manner. The main IGH repertoire results generated during this thesis were peer-reviewed, published in a prestigious journal, are novel and are of general interest to the immunology community.

The dissertation is a 93-page document, structured into five main chapters including an introduction into general B cell biology, B cell receptor repertoires and design of the PhD study (chapter 1), a comprehensive and very well-summarized literature review on adaptive immune responses, B cell differentiation and IGH repertoire sequencing (chapter 2), methodology applied during the thesis (chapter 3), results generated and graphic illustration (chapter 4), and a final chapter with conclusions on the results in this study (chapter 5).

2. *The relevance of the topic of dissertation work to its actual content:*

Prompted by the emergence and further development of high throughput DNA sequencing (HTS) technologies and library preparation methods, adaptive immune receptor repertoire sequencing has emerged as a new area of research. Analysis of human IGH repertoires has provided important insights into the formation and selection of antibodies in different B cell subsets elicited by various pathogens, other immune conditions, and vaccines. The research field has great potential to unravel mechanisms underlying successful or failed immunity, autoimmune diseases, allergy, cancer, and aging.

The topic of this dissertation – the study of IGH repertoires of antigen-experienced B cell subsets in healthy individuals – is thus timely and the actual content of this work has provided a relevant and novel contribution to fundamental knowledge on features and dynamics of these differentiated B cell populations.

3. *The relevance of the methods used in the dissertation:*

To study IGH repertoires of antigen-experienced B cell subsets, Artem used fluorescence-activated cell sorting to isolate memory B cells, plasmablasts, and plasma cells from peripheral blood of healthy individuals based on an appropriate set of cell surface markers. Six study participants were enrolled with sample collection timepoints at enrolment as well as one and twelve months later. Replicate samples were taken at the one and twelve months timepoints to assess variability. B cell subsets were studied exclusively in peripheral blood, which contains only a small fraction of the total B cells in the human body but is the most accessible and thus also the most studied source of B cells. Methods to isolate cell populations from other tissue sites are only starting to emerge and have not been widely applied but present an avenue for future research on this topic. HTS libraries in this thesis were prepared from RNA templates using the 5' rapid amplification of cDNA ends (5' RACE) approach with introduction of unique molecular identifiers (UMIs) allowing for error correction and barcoding/multiplexing of different samples. The 5' RACE approach is a common HTS library preparation strategy and a good choice for Artem's applications and research questions. Using RNA as opposed to genomic DNA has the advantage that isotype information associated with IGH rearrangements is captured, enabling an interesting analysis of the immunoglobulin isotype distribution in different B cell subsets. Introduction of an additional 5' adaptor during the reverse transcription reaction for subsequent amplification of the region likely reduced primer bias compared to other targeted PCR strategies using multiplexed primers. Libraries were sequenced on the Illumina HiSeq 2000/2500 platform, yielding the desired sequencing depth. HTS data were processed through an adequate analysis pipeline including demultiplexing of barcodes, extraction of UMIs, and alignment to V, D, and J germline segments. General repertoire characteristics such as isotype frequencies, somatic hypermutation (SHM) levels, complementarity-determining region (CDR3) length, IGHV gene usage and repertoire similarity metrics were determined. Clonal lineages were defined as groups of clonotypes with the same V segment, CDR3 length and at least 85% similarity in CDR3 nucleotide sequence, which is a clone definition that has been commonly used in similar studies. In summary, for all aspects of the study adequate, state of the art methods have been used to generate the data.

4. *The scientific significance of the results obtained and their compliance with the international level and current state of the art:*

Major findings of this thesis include i) stability of the longitudinal clonal composition of memory B cells in the same individuals over at least one year, ii) inter-individual convergence in memory B cells, likely driven by common pathogen or vaccine exposure and iii) a proposed mechanism of reactivation of memory B cells with additional rounds of affinity maturation and differentiation into plasma cells. Results generated in this study provide relevant information on the characteristics and dynamics of memory and antibody secreting cells in healthy human subjects as well as general and new insights into human B cell biology. Results have been published in a peer-reviewed journal and meet international level and current state of the art requirements.

5. *The relevance of the obtained results to applications (if applicable):*

Repertoire sequencing experiments produce large datasets requiring specialized bioinformatics pipelines to be analyzed effectively. Reliable interpretation of repertoire sequencing data depends upon correct annotation of IGHV genes, using databases of reference sequences of known germline genes, which are currently far from being complete. Inferring novel IGHV sequence alleles (as opposed to somatically mutated versions of known alleles) from repertoire data is therefore challenging.

In his thesis, Artem and his colleagues have developed a novel approach for the genotyping and inference of new allelic variants of V and J genes (called MiStrainer) with several advantages over existing algorithms (such as TIgGER or IgDiscover), including suitability for both hypermutated and non-hypermutated data and good performance at low sequencing depth. The algorithm has the potential to be developed into a public tool for the genotyping and inference of novel allelic V and J gene variants after further improvement and validation.

6. The quality of publications:

During his PhD thesis, Artem has published a first author manuscript in *eLife* and has co-authored a second manuscript in *Frontiers in Immunology*, both peer-reviewed journals with good impact factors (>8 as of 2021). He has published an additional first author paper in the Bulletin of Russian State Medical University. Artem has presented his work in the form of posters at two conferences, the European Academy of Allergy and Clinical Immunology (EAACI) Annual Congress in 2019 and the EAACI Immunology Winter School in 2022.

7. Summary of issues to be addressed before/during the thesis defense:

There are a few minor comments that may be addressed:

- The table of contents should be updated as indicated page numbers in table of contents list do not match actual content on pages and some subsections are not listed in the table.
- Double check paragraph numbering (e.g., subsection 2.5 is listed twice with different headings on page 29 and 31, one heading on page 37 does not have a subsection number).
- Double check some of the references that are listed in the bibliography but not referenced in the text (e.g., R Core Team, Ginestet).
- In the methodology section a short description on programs used to plot the data and perform statistical analyses may be added.
- The isotype distribution and particularly the higher prevalence of IgA compared to IgG in both memory and antibody secreting cells is interesting and somewhat unexpected considering that IgG is the most abundant antibody isotype in the blood. A short discussion on this topic may be added to the conclusions.
- On page 57, the reference to Figures in the text should be double checked (should Fig. 10, 11 actually be Fig. 11,12?).
- Figures 13, 15, and 17 are a little blurry (compared to high resolution of all other figures) and may be exchanged with higher resolution image.
- What is the isotype distribution of public clonotypes (just out of curiosity and in case it is possible to extract this information)?
- The new MiStrainer tool developed during this thesis is an interesting approach for V- and J-gene allele variant inference and may be very useful for the analysis of less studied, non-European populations. What are the plans for this tool – will it be made available to the public after further improvement and validation? This might be mentioned in the conclusions section.
- Some of the text in Chapter 2.5 about the MiStrainer tool is duplicated in Chapter 4.4 and some additional text is also duplicated in 3.5 and 4.4. You may want to choose the better location for these descriptions in only one of the chapters.
- In chapter 5 you may want to add a short subsection discussing lessons learnt (e.g., study design, choice of methods, is there anything you would do differently in retrospect?) and the way forward (potential follow up studies or analyses in the future).

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*