Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Dmitrii Smirnov

PhD Program: Life Sciences

Title of Thesis: Investigation of the role of SIRT6 in the molecular mechanisms of the gene expression regulation, metabolism and aging

Supervisor: Prof. Ekaterina Khrameeva, Skoltech; Prof. Debra Toiber, Ben-Gurion University of the Negev

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

I would like to thank all the jury members for their thorough reviews of my thesis and suggested improvements. Below I provide my replies to all comments and questions along with specifying corresponding changes made in the thesis text.

Prof. Gabriel Leprivier
Thank you for the review!

Prof. Dmitrii Pervouchine
> It seems that the purpose of Abstract and Introduction have been misinterpreted by the defendant: Abstract must be a short summary of the work, while Introduction must be a pedestrian intro into the context of the problem. Here the purpose of these two sections appear to be switched. The Abstract tells the background, while the Introduction merely is an enumeration of thesis chapters. Instead, the Introduction should give the reader a broad understanding where this thesis stands scientifically rather than be a manual on how to use this thesis.
Thank you for pointing this out! I restructured these sections to make the Abstract more concise and to provide more useful information about the scientific relevance of this work in the Introduction section.

> On p.48 brain-specific SIRT6-knockout is mentioned, but the author never mentioned how this knockout was obtained. Is this a cell line or an animal model? Did the author contribute to the generation of this knockout? It would be great to have at least a paragraph on that page that discusses this (maybe it is somewhere else, but I missed it).
Thank you for the comment! The corresponding information regarding the generation of the SIRT6-KO animal and cell models was added to the Methods section in Chapter 4 (page 50). No, unfortunately, I did not contribute to the experimental part of the work, but I mentioned all my colleagues who conducted experiments in the descriptions of methods and also in the 'Author contribution' section.

> Some sections seem to be not very well connected with the others, e.g., switching to diseased old brains in Chapter 5 needs more motivation.

We expanded this Chapter to provide a description of the relevance of SIRT6 to the pathological brain program.

> On p. 65 it is announced that TP73-AS1 is highly expressed in the aging brain. A curious reader wants to ask a question of what is the function of that antisense transcript, and how is it related to the function of the sense gene. I am sure that there is plenty of literature on this topic. Generally, I have an impression that the thesis here and later (especially in Conclusions) becomes a bit cryptic.

Thank you for pointing this out, I provided the information about known TP73-AS1 functions and its relevance to p73 (TP73) in Chapter 2, section 2.3 (page 28).

Prof. Sergey Dmitriev

> It would be helpful to describe the mouse model in a little more fully: how the specificity of Sirt6 knockout in the brain was achieved ("brain" is actually a quite complex organ in terms of histology, so a minimal information about a promoter used for Cre (?) expression, for example, would be useful);

Indeed, the description of the SIRT6-KO models used in the studies was not very clear, so I added more detailed information about the generation of knockout mice and cells (Chapter 4, page 50), mentioning also the Nestin-Cre system we utilized.

> p.3: "We showed that SIRT6 may interact with YY1 and… SIRT3 and SIRT4… to facilitate the transcription of the mitochondria-related genes." – I found no direct evidence for the (direct) interaction of YY1 and sirtuins in this study, except binding to the same promoter regions (which may be quite long) – so it would be good to explain or state this more cautiously.

In fact, the co-binding of SIRT6 and YY1 using co-immunoprecipitation analysis was shown in one of the previous studies published by my colleagues (Stein et al., 2021), where they also provided a number of candidate genes/pathways targeted by the joint regulation of these two proteins. In my thesis, I focused on their cooperation in the context of mitochondria-related regulation, since the very fact of their interaction in the brain has already been shown. In addition, my work mostly relies on bioinformatics methods, therefore this line of analysis was supported by the analysis of publicly available ChIP-seq datasets, that do not reflect interaction mechanisms directly. However, I agree that "interact" might be overly strong in this case, so I slightly rewrote this sentence: "*We showed that SIRT6 may **cooperate** with YY1 (Yin Yang 1) and two other mitochondria-residing sirtuins (SIRT3 and SIRT4) to facilitate the transcription of the mitochondria-related genes*."

> The novel pipeline for untargeted lipidomics data analysis: What is a difference compared to the "old" pipelines? – It should be clarified.

The main advantage of our lipidomics pipeline over the others is that it performs all the steps of data analysis (both upstream and downstream stages), suitable for large lipidomics datasets and has a stable implementation in R. The aforementioned advantages were also listed in Chapter 3, pages 30-31.

> At least in humans, YY1 is one of the major transcription factors regulating the LINE-1 retrotransposon expression. It is also known from the studies by the Gorbunova lab that SIRT6 is critical for LINE-1 suppression. Did you try to analyze the effects of SIRT6 on LINE-1 transcript abundance in your study? It would be also interesting to look at YY1 and SIRT6 binding to LINE-1 promoters in ChIP-Seq data.

Thank you for the interesting question! Yes, SIRT6 deficiency was previously linked to the widespread activation of LINE1 transposable elements in many mouse tissues, including brain. Therefore, one of the lines of our analysis was to examine the changes in the transposable element activity using the obtained *TEtranscripts* tool applied to our RNA-seq profiles. Our analysis revealed no significant differences in the LINE1 expression between SIRT6-KO and WT replicates and we also were not able to reproduce the effect of LINE's activation in other publicly available SIRT6-KO datasets, even though we tried transcriptomic data derived from different tissue and cell-specific models. Given the obtained results, we decided to leave this line of analysis and focus on the differential gene expression findings. I will also address this question more thoroughly in the defense.

> p.69: "Together, this data suggests that YY1 is a major TP73-AS1 regulator." – Please explain: I found no solid evidence in this study supporting the statement that YY1 is indeed the "major" regulator.

In our paper (Mazor et al., 2021), we showed additional pieces of evidence for YY1-dependent TP73-AS1 regulation, including TP73-AS1 expression measurements with both depleted and overexpressed YY1, as well as TP73-AS1 promoter activation by YY1 binding. However, I did not include these results in the thesis, since they were obtained by my colleagues. Thus, this sentence was rewritten to avoid overstatements.

> I would be glad to see the "Conclusions" section in the form of clear theses (points), rather than lengthy text, which would be more appropriate to the "Discussion" section.

Thank you for the suggestion! The conclusion section was modified according to your comment and the main research outcomes were listed with bullet points.

> "M. musculus" and gene names should be italized; nomenclature names of mouse genes should be given as follows: Sirt6.

Done.

> Abbreviations should be spelled out on initial appearance in text (e.g. YY1 (TF Yin Yang 1) is spelled out only in the beginning of the 3rd part of the thesis).

Fixed.

> p68: What is "the 824-promoter activity"?

In this study, we used two regions of the TP73-AS1 promoter sequence, 824 and 1124 bp long. The corresponding information was also introduced in the Methods section in Chapter 5 (page 73).

> References should start with the last names of the authors: this will be much more useful to the readers.

Fixed, thank you for the suggestion!

Prof. Carlos Sebastian

> I miss though a "Objectives" section where the main objectives of the thesis are highlighted and contextualized.

The section "Research objectives" was added to the thesis (pages 14-15).

> Bibliography: the format of this section makes very difficult to find the references cited in the text as they are not numbered and the citation in the bibliography section does not start with the last name of the first author (as it appears in the main text).

Indeed, the bibliography format was not clear enough. I restructured this section so all references follow APA format, starting with the last name of the first author and listed in alphabetic order.

> The role of SIRT6 as a regulator of mitochondrial metabolism (and glucose metabolism, which is intimately linked to mitochondrial metabolism) has been extensively studied in a variety of cell types and tissues. It is very surprising that any of these studies is mentioned and discussed in the current thesis, neither in the introduction nor in the results/discussion sections, despite being very relevant to interpret some of the results obtained in this thesis.

Thank you for noticing! I agree, that the existing results on the implication of SIRT6 in the mitochondrial metabolism were not discussed thoroughly. I added a paragraph highlighting the SIRT6's role in glycolysis and gluconeogenesis to Chapter 2 (pages 21-22) and discussed our results in the context of it in Chapter 4's Discussion (pages 68-69) to remedy this omission.

> One full chapter of this thesis is devoted to the set up and description of a new experimental strategy to analyze, quantify and visualize lipidomics data. However, I couldn't find any experiment in the other two chapters (and specially in the metabolimics experiments on SIRT6-KO cells) where this methodology has been employed, thus making difficult to contextualize this chapter in the overall thesis project.

In fact, we did apply our pipeline to the analysis of the mouse brain lipidomics profiles but initially, I did not include this part in the thesis since we were not able to detect changes due to experimental design limitations (a combination of small sample size and unbalanced design) and these results were never published. Now this analysis is included in Chapter 3 ("3.2.1 Analysis of brSIRT6-KO lipidomics profiles" section, pages 44-46) to demonstrate its application for the analysis of the lipid composition in SIRT6-depleted mouse brains.

> While mice brains have been used for the transcriptomic experiments, only cells with SIRT6-KO have been used for the metabolomics. It would have been better to use the same biological material to integrate both omics experiments to obtain a real multiomic approach.

Thank you for the comment! Indeed, the project was originally designed this way, to obtain multilayer omics profiles (transcriptomic, metabolomic and lipidomics) from mouse brains for the subsequent data integration using bioinformatics methods. However, obtained metabolomics and lipidomics datasets were difficult to analyze jointly with RNA-seq data due to the limitations mentioned in my previous answer. Therefore, we decided to stay with the mESC metabolomics dataset, since it is devoid of these shortcomings and also has higher metabolite coverage, compared to the obtained brain profiles.

Dr. Leonid Peshkin

> The literature review is substantial, but the overall flow and narrative could be improved

Thank you for the suggestion! The "Review of the Literature" section was restructured.

> Chapter 3 should state clearly what was the contribution, which challenges arose in the process of creating the pipeline and how these were addressed. Clearly state the novelty. Also discuss separately the issues of lipid species identification, absolute and relative quantitation and give a clear idea of what can be done confidently and what are the outstanding issues

Thank you for pointing this out! I added sections discussing the advantages of our pipeline (comprehensiveness, applicability for large lipidomics datasets) compared to existing tools and accompanied challenges on pages 30-31 of Chapter 3. The issues associated with lipid quantification and annotation were discussed in the sections "3.3.1 Future Challenges" and "3.1.15 Annotation", respectively.

> Chapter 4 should discuss the limitations of making the inference about activity of such pathways from transcriptomics data as opposed to proteomic and metabolomic measurements and posttranslational controls of enzymes. Separately discuss the relevance of cell culture validation to the observations done in tissue based data

I agree, that such limitations should be discussed in the thesis, so I introduced the corresponding changes in the section "4.3 Discussion".

> Chapter 5 needs to be augmented / expanded it appears to be hastily put together

Thank you for the suggestion, Chapter 5 was expanded with a more detailed preamble and the description of several methods added (pages 72-74).

> Discussion would benefit from some evolutionary perspective on SIRT family expansion and functional conservation, tissue and cell type distributions.

Indeed, the evolutionary, functional and tissue-specific aspects of sirtuin activity were described poorly in the thesis, so I added additional information about it in Chapter 2, discussing the following topics:
- evolutional and functional conservation of sirtuins (page 20)
- tissue-specific functions of SIRT6 (both text and the Figure 2-3, page 25)

Prof. Philipp Khaitovich

> There is a slight omission in the detailed discussion of SIRT proteins link to brain metabolic regulation based on the study results.

Thank you for pointing this out! I expanded the discussion of SIRT6 implication in cellular metabolism in the context of study results (section "4.3 Discussion", pages 68-69).

Prof. Dmitry Ivankov

> In Chapter 3, Dmitry developed a pipeline of processing the lipidomics data coming from LC-MS method. I would like to stress that "Methods" section has a central position in this Chapter since it describes the organization of the pipeline. Taken into account common supporting role of a "Methods" section, it might be worth to entitle it as "Pipeline", "Workflow" or similarly.

Thank you for the suggestion, the title of the corresponding section was changed to "Workflow".