Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Aleksandra Galitsyna

PhD Program: Life Sciences

Title of Thesis: Chromatin folding in individual cells

Supervisor: Professor Mihail Gelfand

I want to acknowledge the high appreciation of my work and constructive comments by all Jury Members. I introduce minor modifications to the text of my Thesis in response to specific concerns listed below.

Prof. Veniamin Fishman

Comment: "Each of the chapters four to eight describes a different research problem addressed by Alexandra. The largest part of the obtained results describes chromatin folding in individual cells, in agreement with the topic of the dissertation. Other studies also belong to the field of 3D-genomics; however, it seems that they are not always directly connected to the former part focused on single-cell genomics. This is not an objection; in contrast, it is commendable that Alexandra was able to explore different aspects of chromatin biology during her Ph.D. studies."

I explain that Chapters 4-6 should be read as prolegomena to the single-cell spatial genomics apex of my thesis. As I highlight in Chapters 4-7, research on bulk 3D-genomics is necessary for single-cell Hi-C study.

Comment: "In Fig. 2-1, it seems that 30-nm fiber is schematically shown. It is now believed that a 30-nm fiber does not exist in the majority of cells under physiological conditions (doi: 10.4161/epi.28297), thus Fig. 2-1 should be revised."

Despite some researchers actively discuss in vivo 30-nm fiber [1], find novel evidence for it [2, 3], and include it as the schematics of chromatin compaction in fundamental reviews on epigenetics [4], I agree that our knowledge about 30-nm fiber limits it to very rare and specific cases not related to Drosophila chromatin. I replace the scheme with one omitting the nucleosome fiber formation.

1. Chen, Ping, Wei Li, and Guohong Li. "Structures and Functions of Chromatin Fibers." Annual Review of Biophysics 50 (2021): 95-116.
2. Risca, Viviana I., et al. "Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping." Nature 541.7636 (2017): 237-241.
3. Swygert, Sarah G., et al. "Chromatin fiber folding represses transcription and loop extrusion in quiescent cells." bioRxiv (2021): 2020-11.
4. Tsou, Pei-Suen, John Varga, and Steven O'Reilly. "Advances in epigenetics in systemic sclerosis: molecular mechanisms and therapeutic potential." Nature Reviews Rheumatology (2021): 1-12.

Comment: "In Fig. 4-1, it is not clear what are white spaces between filled segments. One may speculate that these are inter-TAD regions; however, it is not obvious, because many TAD-callers call TADs as adjustment intervals not allowing inter-TADs. I would suggest explaining what these white

spaces mean explicitly."

I explain that in the caption to the figure.

Prof. Peter Kharchenko

Comment: "One concern is that given multi-author efforts, from the thesis itself it is not entirely clear what aspects of the presented analyses were carried out by Aleksandra."

Prof. Veniamin Fishman

Comment: "I enjoyed this style of presentation; however, sometimes it was difficult to distinguish the author's contribution. For example, chapter five ("A machine learning framework for the prediction of chromatin folding in Drosophila using epigenetic features") is based on the paper where Alexandra is one of the two co-first authors with equal contribution. I was not able to access Alexandra's contribution to this work, neither from the manuscript text nor from the thesis. Note that this comment should be treated as an editorial suggestion rather than an essential concern about the author's contribution, because in the majority of cases, including the part describing single-cell Hi-C data analysis, Alexandra's contribution is clearly evident."

I expand the "Publications" section at the beginning of the Thesis and elaborate on my contribution in Chapters 4, 5, 6.

Prof. Anton Goloborodko:

Comment: "What are the fundamental limits for bias extraction in Hi-C? In other words, can we expect one day to be able to cleanly extract all experimental biases from Ki-C, w/o removing any useful information? What kind of additional experiments could one design to help with this question?"

I expand the discussion of limits for bias extraction in Chapter 6.

Comment: "What would single-cell Hi-C maps look like if we had an "ideal" experiment, where 100% of DNA fragments would get ligated and recovered? Would we be able to tell protein-supported loops from sporadic contacts? What would TADs and compartments look like? What kind of additional information would we be able and not able to extract from such maps? How badly would the presence of sisters and homologs confuse us?"

I elaborate on "ideal" scHi-C in Chapter 8.

Prof. Petr Sergiev:

Comment: "As a future path to the development of better understanding of chromatin folding via HiC based methods I would suggest to address the possibility to map interactions of (i) homologous chromosomes and sister chromatids, (ii) non-unique DNA regions, such as centromeres, telomeres, nucleolar rDNA etc. as chromatin folding of these regions might and even definitely IS different from that of unique DNA regions."

I elaborate more on applications of scHi-C for homologous chromosomes and sister chromatids in Chapter 8.

Prof. Dmitry Ivankov:

Comment: "The only question I have to Aleskandra concerns the analysis accomplished in Chapter 6: How do you think, would it be more convenient and effective to analyze correlation matrix (see Fig.3C of Lieberman Eiden et al., 2009) rather than contact matrix? Maybe, in this scenario one could avoid iterative correction and/or have better characteristics for bulk/single-cell analysis?

I explain the limitations of correlation analysis in Chapter 6.

Comment: “The Glossary looks like a list of abbreviations. Are there any terms that are not abbreviated?”

Non-abbreviations added to the Glossary.

Comment: “Page 12: “Striking conservation of mechanisms was suggested because TADs are present across a wide range of species, including mammals, insects, and nematodes.” May the formation of TADs is inevitable? It may turn out that the existence of long DNA molecules within nuclear is impossible without formation of TADs.“

I add the note that there are specific cases when Hi-C maps have no TADs and the formation of bulk TADs. This problem is not related to TADs formation in stochastic single-cell Hi-C maps already explained in Chapter 7.

Comment: “Page 15: “Proline … is a very rigid aminoacid, typically not forming hydrogen bonds.” This should be reformulated because word “typically” implies that sometimes hydrogen bond still can be formed. However, as we know from physics, if a chemical group can form hydrogen bond then it must always form it in water. What Aleksandra meant here is, probably, that N-terminal group of proline never forms a hydrogen bond, thus, making proline unfavorable in most regions of alpha helices and beta-sheets.”

Thank you, this is an important explanation of my logic. I added it to Chapter 7.

Comment: “The other remarks are just few typos, the list of them is given to Aleksandra.”

Minor remarks are corrected throughout the text.