Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Elena Kurilovich

PhD Program: Life Sciences

Title of Thesis: The role of genome maintenance proteins in primed CRISPR adaptation by the type I-E CRISPR-Cas system

Supervisor: Professor Konstantin Severinov

Name of the Reviewer: Francisco J. Martinez Mojica

I confirm the absence of any conflict of interest	**Date: 04-08-2023**

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

The PhD thesis is very well structured, including Abstract, lists of symbols/abbreviations and figures, Chapter 1 reviewing the literature, Chapter 2 listing the objectives of the research performed, Chapter 3 on materials and methods used, Chapter 4 presenting the results obtained, Chapter 5 discussing the results, Chapter 6 highlighting the main conclusions and a final Chapter 7 with a very complete list of references.

The literature review is accurate, updated and well written, covering the diversity of prokaryotic defense systems, with a main focus on CRISPR-Cas and the best characterized innate immunity systems, the genome maintenance systems and known interactions between their respective components, providing information that is relevant to interpreting and understanding the results obtained in this work.

Materials and Methods chapter is well organized, and information provided would allow for the reproduction of most experiments without reference to other sources, exception made by results presented in section 4.1 (see below). Methods comprise from classical microbiology techniques to the

most advanced molecular biology tools and involve complex sequence data analyses as well as the use of unpublished, original approaches suited to accomplish the present study.

Regarding Chapter 4, it is worth to highlight that the candidate is very careful including only her own results and acknowledging the work done by other researchers involved in the tree published original research articles related to the thesis and coauthored by the candidate:

> 1. Kurilovich, E., Shiriaeva, A., Metlitskaya, A., Morozova, N., Ivancic-Bace, I., Severinov, K., & Savitskaya, E. (2019). Genome Maintenance Proteins Modulate Autoimmunity Mediated Primed Adaptation by the *Escherichia coli* Type I-E CRISPR-Cas System. Genes, 10(11), 872.
>
> 2. Shiriaeva, A; Kuznedelov, K; Fedorov, I; Musharova, O; Khvostikov, T; Tsoy, Y; Kurilovich, E; Smith, G; Semenova, E; Severinov, K (2022). Host nucleases generate prespacers for primed adaptation in the *E. coli* type I-E CRISPR-Cas system. Science Advances, 8(47):eabn8650.
>
> 3. Dimitriu, T; Kurilovich, E; Lapinska, U; Severinov, K; Pagliara, S; Szczelkun, M; Westra, E. (2021) Bacteriostatic antibiotics promote CRISPR-Cas adaptive immunity by enabling increased spacer acquisition. Cell Host & Microbe, 30(1):31-40.e5.

At present, the paper # 1 has received 5 citations in the Web of Science (WOS) Core Collection. The year of its publication (2019), *Genes* got a Journal Impact Factor (JIF) 3.759 and JIF percentile 69.94 within the Genetics & Heredity category, being ranked in Q2 (54/178). Paper #2 has received 1 citation in the WOS Core Collection. The year of the publication (2022), *Science Advances* got a JIF of 13.6 and JIF percentile 91.1 within Multidisciplinary Sciences category, being ranked in Q1 (7/73). Paper #3 has received 12 citations in the WOS Core Collection. In the publication year (2021), *Cell Host & Microbe* got a JIF of 31.316 (JIF percentile 98.72), being ranked in Q1 (1/39) within Parasitology category. The scientific quality of the three papers is noteworthy.

Results shown in section 4.1 of the thesis deal with the analysis of spacer acquisition by the I-F CRISPR-Cas system in *Pseudomonas aeruginosa* in the presence of antibiotics, revealing the promotion of this CRISPR activity by bacteriostatic antimicrobials. As acknowledged by the candidate, this study is not particularly connected to the main objectives of the thesis. Probably that's why materials and methods related to this section are not included in Chapter 3, relaying on their description in publication #3. Moreover, results are discussed at the end of section 4.1 rather than in Chapter 5 and related conclusions are not included in Chapter 6. Section 4.2, on the modulation of type I-E primed adaptation by genome maintenance proteins in *E. coli*, presents results based on the publication #1. Section 4.3, on the participation of host nucleases in the generation of prespacers for primed adaptation by the *E. coli* type I-E CRISPR-Cas system, presents results from the publication #2. Section 4.4 presents unpublished results on the *in vivo* detection by an original method of the half-integrated prespacers during primed adaptation by the *E. coli* type I-E CRISPR-Cas system.

The candidate extensively discusses on the research carried out and clearly discerns among speculations and what is proven after the results obtained. Also, she proposes well substantiated explanations for unexpected or uncertain findings, proposing further experimentation to clarify them.

The biochemical and functional characterization of CRISPR-Cas systems is of great interest to understand the arms-race between viruses and prokaryotes and its biological consequences. Even though the spacer integration step of CRISPR adaptation (i.e., acquisition of CRISPR spacers) has been characterized for a few CRISPR-Cas systems, the mechanism and proteins involved in the generation of spacer precursors are still largely unknown. This thesis discloses the involvement of particular DNA repair pathways, as well as the contribution of specific proteins and activities, in the generation of spacer precursors acquired by the *E. coli* type I-E CRISPR-Cas system during primed adaptation. By delving into the mechanisms and intricacies of primed adaptation, the study opens up new perspectives that can be crucial for

comprehending the broader context of CRISPR-based immunity. Furthermore, the findings presented in the thesis have substantial potential for practical applications based on spacer acquisition. These applications encompass diverse areas such as biosensing, information storage, and gene expression monitoring. However, owing to the fundamental nature of the research conducted, additional potential applications are currently unforeseeable.

In general, the thesis is very interesting, is clearly written, and both organization and presentation are correct. In my opinion, the thesis meets all the requirements for defense as it is.

Below are a few questions and minor formal/grammar recommendations that the candidate may wish to address.

- Define tracrRNA as "trans-activating CRISPR RNA" (list of abbreviations and Page 19)
- Add "Cas" to the list of abbreviations and define it as "CRISPR-associated sequence" (see also Page 15)
- Note that letters size of some figure quotations in the text are different from the rest.
- Page 12. Check the sentence "Major groups the bacterial and archaeal defense systems that may be divided"
  Review the sentence "each of them containing from 1 to **100** CRISPR repeats" considering that CRISPR arrays with several hundred repeats have been reported.
- Be consistent though the text when referring to CRISPR-Cas/CRISPR systems (see for example first line on page 16).
- Page 17. Check the word "**bi**undergoing"
- Page 27. Check the sentence "A similar construct with 5' overhangs **are** not bound"
- Page 28. Replace "(3'-CTT-5')" with either "(**5'**-CTT-**3'**)" or "(3'-**TTC**-5')"
- Pages 28, 95. It is correctly stated that spacer integration starts "with the Cas1-catalyzed nucleophilic attack by the 3'-OH of the prespacer at the phosphodiester bond between the leader and the first repeat in the CRISPR array, leading to formation of a half-site product (Figure 3) [149]". However, Figure 3 shows, and it is claimed in reference 149, that the leader-repeat junction is the second, rather than the first, target site for integration. An appropriate figure/reference (i.e. doi:10.1126/science.aao0679; doi:10.7554/eLife.08716) should be used/cited instead.
- Page 31. In agreement with the general thought, it is stated that "RecBCD binds blunt or nearly blunt dsDNA ends [160] and starts unwinding and degrading both strands". However, it has been proposed (doi: 10.1128/mmbr.05026-11; doi: 10.1016/bs.adgen.2022.06.001) that RecBCD does not degrade DNA but translocate on the DNA from the DSB and when reaches a Chi site, nicks it. This possibility might be discussed by the candidate.
- Page 36. Replace "*Streptococcus thermophil**es***" with "*Streptococcus thermophil**us***" and "*P. aeruginosa **PA14***" with "*P. aeruginosa* **PA1**4"
- Page 38. Replace "1.7.2. Clinical relevance" with "1.7.2. Clinical relevance **of *P. aeruginosa***"
- Section 4.1. This section, along with related information in other chapters (i.e., Chapter 1), could be removed, as this study is not particularly connected to the main objectives of the thesis. Alternatively, I suggest to either add related materials and methods to Chapter 3 or provide further details in section 4.1 to ensure sufficient understanding of the experiments performed.
- Fig. 15. Replace "*Δrec **ΔB**recJ"* with "*Δrec**B** **Δ**recJ"*
- Section 4.2.1. The heading of this section, "Primed CRISPR adaptation is impaired in *ΔrecJ, ΔrecB ΔrecJ* and *ΔrecB ΔsbcD* mutants", does not exactly correspond to its content. Even though figures in this section show data with additional single (*recJ, sbcD, sbcB*) and double mutants, only results with single

recB, recC and *recD* mutants are outlined in the main text. The results obtained with the other mutants should be mentioned and discussed as well.

- Page 65.
    - "induction of self-targeting in the mutants caused SOS response, as judged by cessation of culture growth (Figure 15),…" Indeed, SOS response could be, and probably is, at least in part responsible for this phenotype. However, without additional experimental prove (use of appropriate controls, such as inactive *cas3* mutants, or demonstration of SOS induction with SOS response reporters), the contribution of protein overexpression cannot be dismissed and should be considered as massive protein expression after transcription induction might account for the growth delay, CFU decrease and cell elongation (shown in Fig 17, but not mentioned in the text) observed, independently of interference activity.
    - "the addition of *cas* genes inducers as judged by differential staining of live and dead cells **(Figure 17)**". Refer to Figure 18B as well.
- Fig. 23. Define "*WT* (-Ind)"
- Fig. 24. Replace "Fragments mapping to the NT-strand upstream of the PPS and to the T-strand downstream of the PPS are shown in **green**" with "Fragments mapping to the NT-strand upstream of the PPS and to the T-strand downstream of the PPS are shown in **gra**y".
- Page 80. Add ")" to the end of the sentence "… presumably generated by random DNA fragmentation"
- Page 95.
    - Replace "Th**a**n the second" with "Th**e**n**,** the second"
    - "A new method developed in our lab". A schematic figure illustrating the HIPs detection method would be appreciated.
- Page 97. Replace "RecJ **of** ExoVII" with "RecJ **or** ExoVII"
- Page 100. Clarify if the sentence "No preferences in selection of new spacers with specific orientation of the PAM sequence have been reported so far" refers to just naïve adaptation.
- Page 101. Review the sentence "… for plasmids become unstable in Rec mutants"
- Discussion. Your model of primed adaptation seems to apply to the acquisition of spacers from both the upstream and the downstream regions of the PPS. If so, do you assume that after PPS binding the PAC translocates either in one direction or the other from the PPS? Alternatively, do you think it is possible that upstream prespacers are generated by the PAC and downstream by other proteins/complexes?
- Bibliography. A few references are incomplete or contain errors, such as:
    - The month of publication is missing (e.g., pp. 466–472, 2005; 467–469, 1992)
    - Additional numbers appear before the volume (e.g., Science (80-.); Nat. Struct. Mol. Biol. 2004 119; Nat. Struct. Mol. Biol. 2011 186)
    - Author's initials are duplicated (e.g., D. C. D. C. Swarts; S. J. J. S. J. J. Brouns; S. J. J. J. Brouns; P. M. P. M. Nussenzweig)
    - Not abbreviated (e.g., Nature Reviews Microbiology; Methods in Molecular Biology) or incomplete (e.g., Proc. Natl. Acad. Sci.,) journal names
    - Doi is included (e.g., http://dx.doi.org/10.4161/rna.8.3.15190)

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*