CHARLES UNIVERSITY
Faculty of science

In Prague, 22nd of February 2020

Review of the PhD thesis of Anna Nenarokova

The PhD thesis submitted by Anna Nenarokova sums up Anna's research on the genome analyses of *Blasthocrithidia*, specifically focused on the re-assignment of the stop codons to amino acid coding codons.

The thesis contains an introduction on *Blasthocrithidia* followed by description of gene expression in kinetoplastids. Final part of the introduction is dedicated to the evolution and derivations of the genetic code in eukaryotes.

This section is nicely written and summarizes necessary information but remains largely on the surface of quite complex already known facts on kinetoplastid gene expression and its regulation. The thesis then contains method and result sections, which describe the actual findings related to re-wiring of the stop codons.

There is a short summary section and conclusions at the end of this part on yet unpublished data. This work actually follows on the original discovery of stop codon re-assignment in *Blasthocrithidia* and attempts to understand its functional consequences and evolutionary history. It is obvious from the work that a lot has been done but as we usually have to admit: "we still don't understand the mechanism". There is some strange wording in the introduction and there some formal shortages such the lack of the legends for the figures but the overall quality is high.

Indeed, it is a complicated matter and I would rather recommend to dedicate the thesis to the project, which Anna actually successfully completed and published as a first author in mBio. The *Blasthocrithidia* project could be added as an independent manuscript. Hence, in the current shape, it is not clear whether the introduction relates to the entire thesis or just to one of the particular projects Anna was involved in.

I have following specific questions:

1. **There is a following reasoning for the universality of the genetic code mentioned in the thesis: „the existence of multiple genetic codes makes horizontal gene transfer almost impossible" – Does it also stand for eukaryotes? How frequent is the HGT among eukaryotes that the universality of the genetic code is supposed to be one of the criteria for their "successful evolution"?**
2. **Do you see less HGT in species with the re-assigned (stop) codons?**

In total, Anna has co-authored 8 publications, which are part of the thesis. All the publications went through the rigorous review process, so there is little room left for the thesis reviewer. In general, it is impressive amount of work, which deserves my congratulations. On the other hand, I see a job of bioinformaticians like Anna quite unfair as they are less frequently given a lead role in the project and rather

FACULTY OF SCIENCE
Pavel Doležal,Ph.D.

address: Albertov 6, 128 00 Praha 2 **phone:** +420 325 873 909 **e-mail:** pavel.dolezal@natur.cuni.cz

web: www.natur.cuni.cz

assist others with data analysis. Hence, in addition to *Blasthocrithidia* project, I find the publication in mBio as essential contribution to the whole thesis.

The publication "Causes and Effects of Loss of Classical Nonhomologous End Joining Pathway in Parasitic Eukaryotes" describes interesting phenomenon of the loss of NHEJ pathway in parasitic protists. Anna and co-authors also hypothesize why this could happen and if it brings any benefits to the parasite.

I have following specific questions:

- **What is the minimal set of c-NHEJ components representing the functional pathway?**
- **Are there parasitic protists genetically manipulated towards the loss of NHEJ so that they rely preferentially on homologous recombination?**
- **And vice versa, as a *Giardia* researcher I would be very happy to introduce NHEJ to this organism, what should I do?**
- **How can you experimentally test (forget the presence or absence of the executing protein components) whether c-NHEJ or other types of NHEJ are present in a particular organism?**

To sum up, the PhD thesis of Anna Nenarokova represents very solid and exciting piece of scientific work. Due her obvious professional qualities Anna has been involved in more less unrelated large scale genomic and proteomic studies, which all resulted into high quality papers. The main body of the thesis summarizes her experimental efforts to understand the molecular background of stop codon re-assignments in *Blasthocrithidia*. While yet unpublished and not completed, this part represents an interesting insight into an intriguing biological phenomenon, which questions some of the basics of the molecular biology. Importantly, Anna also succeeded in publishing an interesting study on the evolution of NHEJ pathway in parasitic protists as a first author.
Thus, I am entirely confident that the quality of the submitted thesis is suitable for its defence and awarding a doctoral degree to Anna.

Pavel Doležal, Ph.D.

FACULTY OF SCIENCE
Pavel Doležal,Ph.D.

address: Albertov 6, 128 00 Praha 2 **phone:** +420 325 873 909 **e-mail:** pavel.dolezal@natur.cuni.cz
web: www.natur.cuni.cz

INSTITUTE OF MICROBIOLOGY
of the Czech Academy of Sciences

Dr. Miroslav Oborník
Head of the Committee for PhD studies
in Molecular and Cell Biology and Genetics
University of South Bohemia
Faculty of Science
České Budějovice
Czech Republic

Prague, Jan 31, 2020

Evaluation Report for the Ph.D. award

Name of the Student: **MSc. Anna Nenarokova**

Title of Thesis: **Genomics of Blastocrithidia, a trypanosomatid with all three stop codons reassigned.**

Name of the Supervisor: **Prof. RNDr. Julius Lukeš, Csc.**

Dear chair, dear committee members,

Here I submit to your hands the evaluation report on the dissertation work of MSc. Anna Nenarokova for the award of Degree of Doctor of Philosophy (Ph.D.).

MSc. Nenarokova's work is based on 8 already published papers, one of which features MSc. Nenarokova as the solo first author (MBio).

As the title states, her work describes genomic studies of Blastocrithidia, the trypanosomatid where all three stop codons have been reassigned to encode specific amino acids. In particular, this work suggests that UGA encodes tryptophan, while UAG and UAA encode glutamate. In addition, it was proposed that UAG and UAA are decoded by the Blastocrithidia-specific tRNAs with the mutated CUA and UUA anticodons, originating from the canonical glutamate decoding tRNAs, while UGA in the cytosol is likely decoded by the tryptophane tRNA with the canonical CCA anticodon, however, bearing specific mutations in the anticodon stem loop. One of the major conclusions is that contrary to a previous report, only UAA acts as a

Videnska 1083
142 20 Prague 4
the Czech Republic

tel. +420 241 062 288
+420 605 863 709
fax +420 241 062 501

e-mail: valasekl@biomed.cas.cz
https://mbucas.cz/en/

genuine stop codon that is able to trigger translation termination, however, only in the context-dependent manner, determined by the markedly low GC-content of 3' UTRs in this specie. The other major conclusion has that several features of the Blastocrithidia genomes related to the genetic code reassignment can be explained by a simple assumption, which is that the common ancestor of the genus Blastocrithidia experienced a significant period of the AT mutational pressure.

MSc. Nenarokova's doctoral thesis is presented in a short format with a classical organization for such a format. It is written in very good English and – with a few exceptions listed below - provides the reader with a broad overview of the entire problematic, clear description of all unpublished results combined with their discussion, followed by a comprehensive and up-to-date list of references, captured with the reprints of all 8 publications.

Below I list several minor issues that I noticed followed by my specific questions:

- Description of "translation" on page 19 is very poor, focused on initiation factors 4G, 4E and PABPs only, which is meaningless in this form; I understand that the author wished to introduce PABPs that become important later on, however, in this way it was not intuitive and did not help the reader at all.
- Description of translation termination in eukaryotes (especially the role of eRF3) is, according to my knowledge, incorrect; please explain the committee the most up-to-date model of termination/recycling in eukaryotes.
- The figures are mislabeled from page 33 on, which is confusing – I suggest to correct it in all printed versions manually.
- Some refs in the text lack brackets, e.g. Laird, 1959, Doflein, 1951.

- "Curiously, UAG and UAA are usually reassigned together to encode the same amino acid, while UGA is reassigned independently." Any idea why UAG and UAA are reassigned together?
- Could the significant differences in the GC-content and/or the loss of the NMD pathway component(s) among organisms from the same clade or sister clades serve as a potential predictor of genetic code reassignments?
- "Moreover, the truncated version of the human PABP protein is even less selective – it binds AT-rich sequences only 2 times less efficiently than the full version (Sladic et al., 2004)." What about the GC-rich sequences? Is it still able to discriminate sharply between AU-rich vs. GC-reach content? Is it possible to examine these discrimination abilities of PABP1 of Blastocr. vs. Jaculum vs. other trypanosomatids biochemically to prove or disprove this theory? If yes, please suggest how you would tackle it.
- Can you create a reporter with the UAA stop codon followed by the GC-rich 3' UTR (poor stop) vs. AU-rich 3' UTR (strong stop) with a single in-frame UAA stop codon further downstream followed by the AU-rich sequence to produce two distinguishable protein products of a different length to support your interesting in silico observation? If yes, please design such an experiment.

- Page 49; this is very interesting because we have data showing that the identity of bases 28 and 42 forming the 4th pair (out of five) of the anticodon stem loop of tRNAGlu critically determines the efficiency of the UAG stop codon readthrough in the budding yeast ...
- How would you define a source of the "AT-mutational pressure"?
- As for your evolutionary scenario presented on page 54 ... if I understand it correctly, the major trigger/facilitator of these reassignments was the deactivation of the NMD pathway. If so, your retrospective look should start with the loss of NMD, allowing the reassignment of UGA as the first step, and so on...

Taken together, this thesis represents a rather large amount of the quality work of this PhD candidate and clearly demonstrates her experimental, as well as intellectual skills that are required to obtain a Ph.D. degree. Therefore, I gladly recommend acceptance of this PhD. work and congratulate the author on such a nice piece of work.

Res. Prof. Leoš Shivaya Valášek, Ph.D.
Laboratory of Regulation of Gene Expression
Institute of Microbiology of the Czech Academy of Sciences
Videnska 1083, 142 20, Prague 4
The Czech Republic