Přírodovědecká fakulta
Faculty of Science

Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

OPPONENT'S REVIEW ON BACHELOR/DIPLOMA* THESIS

Name of the student: Alexandra Kogler

Thesis title: Alexandra Characterization of Bartonella species infecting North American Triatominae (kissing bugs)
Supervisor: Eva Nováková

Referee: Ivan Fiala
Referee`s affiliation: Institute of Parasitology, BC CAS

	Point scale[1]	Points
(1) FORMAL REQUIREMENTS		
Extent of the thesis (for bachelor theses min. 18 pages, for masters theses min. 25 pages), **balanced length of the thesis parts** (recommended length of the theoretical part is max. 1/3 of the total length), **logical structure of the thesis**	0-3	3
Quality of the theoretical part (review) (number and relevancy of the references, recency of the references)	0-3	3
Accuracy in citing of the references (presence of uncited sources, uniform style of the references, use of correct journal titles and abbreviations)	0-3	3
Graphic layout of the text and of the figures/tables	0-3	1
Quality of the annotation	0-3	3
Language and stylistics, complying with the valid terminology	0-3	2
Accuracy and completeness of figures/tables legends (clarity without reading the rest of the text, explanation of the symbols and labeling, indication of the units)	0-3	1
Formal requirements – points in total		16
(2) PRACTICAL REQUIREMENTS		
Clarity and fulfillment of the aims	0-3	3
Ability to understand the results, their interpretation, and clarity of the results, discussion, and conclusions	0-3	1
Discussion quality – interpretation of the results and their discussion with the literature (absence of discussion with the literature is not acceptable)	0-3	1
Logic in the course of the experimental work	0-3	2

* Choose one

[1] Mark as: 0-unsatisfactory, 1-satisfactory, 2-average, 3-excellent.

Completeness of the description of the used techniques	0-3	2
Experimental difficulty of the thesis, independence in experimental work	0-3	2
Quality of experimental data presentation	0-3	1
The use of up-to-date techniques	0-3	2
Contribution of the thesis to the knowledge in the field and possibility to publish the results (after eventual supplementary experiments)	0-3	1
Practical requirements – points in total		15
POINTS IN TOTAL (MAX/AWARDED)	**48**	**31**

Comments of the reviewer on the student and the thesis:

Main goal was to evaluate phylogenetic diversity of *Bartonella* species and their prevalence in *Triatoma rubida* from two sampling sites in USA sampled in 2018 and 2019. I appreciate that multilocus sequencing strategy was used for the thesis plus comparing results with already obtained 16S amplicon sequencing. I liked the M&M section with all relevant information to the topic.

The problematic issue of the thesis was obtaining good quality sequences from direct sequencing of PCR products. Only a fraction of the obtained PCR products gave good quality sequence that could be used for phylogenetic analysis, which is essential for fulfilling the main aim of the thesis. See my questions below concerning this problematic.

Considering that the phylogenetic diversity was the main goal of the thesis I was wondering why only ML analysis was performed. I would recommend comparing the results of at least one another method of phylogenetic reconstruction, such as Bayesian inference, as usually is done for the phylogenetic studies. Moreover, one of the issues discussed in the thesis was the low bootstrap support of ML – therefore I would expect that this analysis would be based on the proper bootstrap replicate sampling and not only on very low number of 100 replicates. My other concern is with the quality of the graphic outcome of the phylogenetic trees. In M&M, there is a note that trees were graphically edited in Inkscape, however, the graphical outcome is of very very poor quality with no graphical edition (even with not complete taxon names in many cases and lacking of italics in case of Latin names). I understand that explanation of the phylogenetic relationships is sometimes difficult in the text form but here I feel that it was done very superficially with some missing/erroneous information, for example comments of the analysis of rpo15 (page 33 and Fig. 15): It is said that there are two main branches *B. clarridgeiae* and *B. vinsonii*. However, obtained samples are more closely related to *B. rudakov*… (here is incomplete species name) and unnamed *Bartonella* sp. rather than to *B. clarridgeiae*. Similarly, main branch *B. vinsonii* contains other two *Bartonella* species clustering close to newly obtained sequences. If the trees would be graphically edited, they can be much more self-explanatory.

In my opinion, discussion is too brief. For example, I was expecting some better explanation of the main problematics of unspecificity of *Bartonella* primers – the primers that were used were published more than 10 years ago so some similar works should exist using these primers.

Although, I feel that from this study more relevant data could be obtained, I suppose that it is a good background for further, more deep studies, that can explain interesting problematics of microbial diversity in Triatominae.

Suggestions and questions, to which the student has to answer during the defense. Mistakes, which the students should avoid in the future:

Fig. 6 shows that really very high percentage of obtained 16S rRNA sequences were of bad quality with mixed signals. As I understand it well, these mixed infections were recognized as negative. What method(s) do you think could help to recognise the mixed signals from the PCR product sequencing? If there are some, why did you not try it at least for some small number of samples? What sequences do you think that might be in the mixed signal when using genus (*Bartonella*) specific primers?

What was the criterion for selection of outgroup species for your four phylogenetic analyses? Why only ML method was used for the phylogenetic analysis and not e.g. Bayesian inference to compare the results from other method of phylogenetic reconstruction (especially if the analysis gave unresolved topologies)?

I was confused with the presentation of the data in Table 5 and 6: there are results for 16S rRNA amplicon sequencing and MLST of four housekeeping genes (including 16S rRNA), but there are four columns in the tables - one for every gene. How can I recognise results from amplicon seq of 16SrRNA and sanger (MLST) seq of 16S rRNA? Legend says that results written in red are indicating high quality sequencing for both approaches. If both approaches mean amplicon and sanger sequencing, why there are red in all four columns if only 16S was done by amplicon sequencing? Or do I understand it wrongly? Species names should be in italics in the tables and in table 6, there is a missing first row with genes indication. Legends also lack the information how the species, listed in the tables, were matched with the obtained sequences. Was it done by blasting?

Text commenting the results of phylogenetic analysis contains always information that e.g. "26 samples and 19 genera were used in the analysis of gltA". What do you mean with the "genera"? – I could see sequences representing members of only single genus in the resulted tree.

Conclusion:

In conclusion, I

recommend

the thesis for the defense and I suggest the grade Very good.[2]

In **České Budějovice** date 10th September 2021

[redacted]

[2] You can suggest a grade, which can be modified during the defe[redacted]wever, if the reviewer is not present at the defense, the grade will not be counted. Grades: excellent (1), very good (2), Good (3), Unsatisfactory/failed (4).