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: Jessica Drozdz

Thesis title: Methods for studying gut parasites and their interaction with the host and the host microbiome
Supervisor: Mgr. Martin Kolísko, PhD.
Co-Supervisor: MSc. Petr Táborský

Referee: Mgr. Hana Tykalová
Referee's affiliation: **University of South Bohemia in Ceske Budejovice, Faculty of Science AND Biology Centre, Czech Academy of Sciences, Institute of Parasitology**

	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	2.5
Quality of the theoretical part (review) (number and relevancy of the references, recency of the references)	0-3	2.5
Accuracy in citing of the references (presence of uncited sources, uniform style of the references, use of correct journal titles and abbreviations)	0-3	1.5
Graphic layout of the text and of the figures/tables	0-3	3
Quality of the annotation	0-3	2.5
Language and stylistics, complying with the valid terminology	0-3	3
Accuracy and completeness of figures/tables legends (clarity without reading the rest of the text, explanation of the symbols and labelling, indication of the units)	0-3	1.5
Formal requirements – points in total	**21**	16.5
(2) PRACTICAL REQUIREMENTS		
Clarity and fulfilment of the aims	0-3	2
Ability to understand the results, their interpretation, and clarity of the results, discussion, and conclusions	0-3	1.5
Discussion quality – interpretation of the results and their discussion with the literature (absence of discussion with the literature is not acceptable)	0-3	1.5

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

Logic in the course of the experimental work	0-3	3
Completeness of the description of the used techniques	0-3	1
Experimental difficulty of the thesis, independence in experimental work	0-3	3
Quality of experimental data presentation	0-3	1.5
The use of up-to-date techniques	0-3	3
Contribution of the thesis to the knowledge in the field and possibility to publish the results (after eventual supplementary experiments)	0-3	2
Practical requirements – points in total	**27**	18.5
POINTS IN TOTAL (MAX/AWARDED)	48	35

Comments of the reviewer on the student and the thesis:

The Bachelor´s thesis by Jessica Drozdz entitled "Methods for studying gut parasites and their interaction with the host and the host microbiome" tries to solve the complicated issue of the RNA isolation from the intestinal contents to gain quality-enough sample for subsequent next-generation sequencing analysis. This procedure should serve as a novel tool in the study of *Blastocystis hominis* biology and should enable the determination of gene expression profiles of this intestinal symbiont/parasite.

The thesis is divided into standard sections and is presented with a high-quality scientific English with a minimum of grammatical or stylistic issues. An Introduction spans 15 pages, Aims of the work are defined on one page, Methods are covered on 4 pages followed by 3 pages of Results, 2 pages of Discussion and a brief Conclusion. On a total of 37 pages, Ms Drozdz cites more than a sufficient number of relevant sources (135). As a whole, the Bachelor´s thesis meets all the formal requirements.

An Introduction is logically structured, well written and defines various morphological forms of *B. hominis*, deals with its life cycle and analyses its putative harmful and beneficial roles in the human gastrointestinal tract. In the context of the thesis, the chapter of "Cysteine proteases" seems redundant. It could probably be concepted in more general way – e.g. Blastocystis-host interaction. Also, in this chapter, Ms. Drozdz confuses/mixes host and *Blastocystis* derived factors, which makes the chapter hard to grasp. As *B. hominis* belongs to the group of unorthodox Eukaryotes, for somebody who is not experienced in the field (as myself), the *Blastocystis* genome is poorly characterized – in a single short paragraph. Especially, when the thesis is directly focused on sequencing and assembly of *B.h.* transcriptome, I would expect more background in this respect. Some relevant work could have been reviewed, such as Denoeud et al., 2011 (doi: 10.1186/gb-2011-12-3-r29). As a minor formal issue, I see the improper introduction of abbreviations in the text. The List of Abbreviations is not complete (ER, SSu rDNA, SEM, IBD, IκBs, NF-κB, AIDS, HIV are missing). Some shortcuts are not explained in the text (DIC, TEM, IL-8) and some are used prior explanation (CB). Introduction of shortcut is reasonable when used more than two times in the text (e.g. unnecessary for SSu rDNA).

Methods section of the thesis is chaotically organized with substantial shortcomings in the description of methods and overall the procedures are not easy to follow. Description of rat infection and generation of own samples for analysis is made partially in the "Cultures" section (3.1) partially in the "Gut samples" section (3.3) and partially in "Collection and storage of the

samples" (3.4). The formulation of cultivation media for *Blastocystis in vitro* cultivation is described in the "Cultivation" section (3.2), however, the procedure *per se* is described in the previous section "Cultures" (3.1) and the reason for doing so we find in the section "RNA extraction" (3.5). Section titles are not chosen appropriately. For example "Gut samples" were not the targets intended for isolation, but its content with the *Blastocystis* cells. Or "Processing of data" it too vague designation. The sequencing procedure is partially described in RNA isolation section.

Further issues in Methods just briefly:
Section 3.1: I miss information about sex and age of experimental animals and the frequency of *Blastocystis* culture passaging.
Section 3.2: A proper way of presenting buffer formulation states the molarity of compound, not the weight of the compound in the volume and different mixing ratios (page 17).
Section 3.3: After the statement: "... caecum, which has previously been shown to contain the large amount of active *Blastocystis* cells." I miss the reference.
Section 3.5: Student states: "..we have used Chloroform and BCP for protein extraction", however it is not clear for which purpose this would be done when the work operates with no protein samples.

From the description of isolation: "add 0.2 ml of BCP/Chloroform per 1 ml of Trizol..." is not possible to deduce what solution was added. BCP, chloroform or their mixture? If BCP was used, then the volume should be 0.1 ml.

RPM values without the rotor radius are not a correct reference of centrifugation speed, use centrifugation speed expressed in g instead.

TRIzol product number and manufacturer is missing.

Two RNA isolation techniques are described almost identically without being obvious when and how was the commercial kit implemented. A brief description of the isolation protocol with the kit should be provided. It is not clear how much of the sample was loaded on the column, whether the procedure and kit buffer volumes were adjusted accordingly and what was the final elution volume. The description of the methods should be accurate enough to enable repetition by somebody else.

Bullet introduction in RNA isolation methods 1 and 2 is not uniform.
Section 3.6: I assume that the bioinformatic processing and analysis of NGS data was performed by the student. Then this fact could have been highlighted in the description.

Was poly-A selection performed?

The last sentence of Methods belongs rather to the Results section (or both).

In the Results section out of four *B. hominis* infected rats, RNA samples were isolated using two evaluated methods. Ms Drozdz was able to obtain 2 RNA samples for further NGS analysis with one tested method only. From the presented description it is not clear whether any kind of optimization was performed to tackle the issue of problematic RNA isolation. No data or figures on RNA quality or integrity were presented, only concentration values for 2 successfully isolated samples are stated. Based on that, the reader can only speculate why "samples showed unreliable results and failed repeatedly for some samples" or why "Qiagen RNA extraction kit failed completely".
It is unsure why there is a description of cultivation of *B. hominis* in the Methods section when there is no further reference of it in the Results. As you state on p. 18 of methods: "The *in vitro* cultures from the stool samples were used as a pre-evaluation test to confirm *Blastocystis* infection and were not important for further investigation." Then where is the need for their inclusion in the Method section then?

Further, Ms Drozdz summarizes the data obtained from Illumina sequencing and subsequent bioinformatic analyses in 4 statistical tables. Table legends are however too concise to introduce the data without further need for searching in the results. In the text itself, I would also welcome, if the description was more detailed and explanatory. I had to re-read several times and go back to methods to understand what the results mean and why certain analyses were performed. But that might simply be my inexperience with such type of data. Somehow, I can not help the impression, that after the problems faced with isolation, when succeeding with the sequencing, more could have been mined from the results obtained. With the ambition to be a pioneering study for future *Blastocystis* RNA expression patterns retrieval and rat gut microbiome alteration due to infection study (as outlined in Aims), neither specific/candidate *Blastocystis* genes are mapped nor analysis of other species hits is performed. Also, the presence of negative control would help in the analysis and evaluation of the data.

In the brief discussion, I missed the detailed evaluation of possible causes for failure and inconsistencies during isolation procedures. It is not clear whether the student is aware of all the possible limitations of the chosen methodology. No reference for the comparison with the similar sample type in published studies was provided. The same objection, lack of comparison with the literature, goes to the discussion of transcriptome data analysis.

In the reference section, there are many inconsistencies in the format of references, especially in the author names. Authors are cited with initials (eg. refs. 4, 13, 15), full names (refs. 2, 8, 9, 10, 14) and various combinations of both (refs. 1, 3, 27) randomly and I was unable to figure out what was the intended citing style. Some references have capitalized author names or journals (5, 22, 49, 93, 101, 102). Capitalization of journal names is random as well. However, the list is complete, and the numbering style was implemented well.

To sum up. As a reader, I see a striking imbalance between the quality of well-written Introduction and disorganized Methods and far too concise Results and Discussion sections. Ms Drozdz adopted several basic and also advanced methods in the study of *B.hominis* biology, and was able to successfully implement them and retrieve usable NGS data. As a profound achievement, I appreciate the own performance of bioinformatic analysis on NGS data by the student. But the data presentation, analysis and discussion are feeble and did not employ the full potential of the data.

Mistakes, which the students should avoid in the future:

A proper way of citing:

Of non-English publication (ref. 106) is with stating the original language in parentheses at the end of the reference, e.g.:

Cimerman, S., M. C. Ladeira, and W. A. Iuliano.2003. Blastocystosis: nitazoxanide as a new therapeutic option. Revista da Sociedade Brasileira de Medicina Tropical 36:415–417. (In Portuguese.)

Electronical source (ref. 127) is with full link and date when accessed e.g.:

Andrews S. FASTQC. A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 29 September 2014.

In vivo is written without hyphen.

Suggestions and questions, to which the student has to answer during the defense.

1. What is the epidemiological situation of *Blastocystis* in the Czech Republic?
2. How was Figure 8 modified from the original?
3. What are the main characteristics of the *B. hominis* genome?
4. Could you explain the principle of RNA extraction with TRIzol reagent? What was the volume of sample used for RNA isolation in both methods? What was the sample/ TRIzol ratio? What does the BCP stand for? How much of the sample was loaded on the Qiagen column and what was the elution volume? What was the binding capacity of the column in the kit?
5. How was evaluated the quality of RNA prior sequencing? Was the integrity of the RNA checked somehow? How was the degradation of the sample by RNases prevented?
6. In which step and how was the RNA extraction kit implemented? What is the rationale for doing so?
7. Can you discuss, what was, in your opinion, the cause of inconsistency of performance or failure of tested RNA isolation methods?
8. What was the purpose of mapping of sequenced reads against sets of transcripts of *Bastocystis* and *Rattus*? The discrepancy between the number of reads mapping to the *Blastocystis* (≈85%) and host RNA (≈1%) is striking. Can you discuss this finding and compare it with the available literature?
9. From the data, were you able to retrieve any information on the microbiome of the rat gut? Is the selected approach suitable for such analysis and why?
10. What other methods except for NGS can be used for gene expression quantification of *Blastocystis?* Compare their pros and cons and evaluate their suitability for the intestinal content analysis.

Conclusion:

In conclusion, I

recommend/ ~~do not recommend~~*

the thesis for the defence and the grade will be suggested after defence.

In **České Budějovice** 14th of July, 2020

signature