

## Jury Member Report – Doctor of Philosophy thesis.

**Name of Candidate:** Song Guo

**PhD Program:** Life Sciences

**Title of Thesis:** Using RNA expression as a quantitative molecular phenotype to study human and vertebrate evolution

**Supervisor:** Professor Philipp Khaitovich

**Name of the Reviewer:** Shuhua Xu

I confirm the absence of any conflict of interest

(Alternatively, Reviewer can formulate a possible conflict)

**Date: 06-11-2021**

*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**

Reviewers report should contain the following items:

- Brief evaluation of the thesis quality and overall structure of the dissertation.
- The relevance of the topic of dissertation work to its actual content
- The relevance of the methods used in the dissertation
- The scientific significance of the results obtained and their compliance with the international level and current state of the art
- The relevance of the obtained results to applications (if applicable)
- The quality of publications

The summary of issues to be addressed before/during the thesis defense

The web system responds very slow in displaying the full thesis, but I made it and read the majority of the text as well as most of the figures and tables.

The thesis title sounds to me is a too big topic, is there some way to make it more specific?

I understand the RNA data are the focus of the thesis in terms of computational analysis and view angle of the questions, however, I would be happy to read some more discussions on the impact or influence of genetic variations. I believe it is applicable given the DNA-level data are also available for some tissues studied in the thesis.

Since the sample size is often not large for the data used in the thesis, which I understand the situation, some evaluation of the statistical power for different sample size might be helpful.

Is there any batch effects in the data given they might be generated from different batches, technical platforms, with different reagent, or different time?

The thesis is overall well-written, but there is still some room to improve the language. I would suggest these issues can be addressed after the candidate do oral defense.

#### 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*