**Examiner’s Report**

**Student:** Yutong Li

**Dissertation / Research Project:** [Title]

**Recommended grade:**

Please indicate % mark in the relevant box below. Grade level descriptors are provided at the end of this document to assist you in determining an appropriate grade.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Class of Honours** | 1st | | | 2nd div I | 2nd div II | | None | | | Fail |
| Grade | A+ | A | A- | B+ | B | B- | C+ | C | C- | D |
| GPA | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| % mark range | 100-90 | 89-85 | 84-80 | 79-75 | 74-70 | 69-65 | 64-60 | 59-55 | 54-50 | <50 |
| **Recommended % mark** |  |  | X |  |  |  |  |  |  |  |

**Please comment on the dissertation under the following headings:**

Introduction & statement of problem

Capable introduction, well-aimed at the research questions. Some of the discussion of Bayesian phylogenetics is a bit of a grab-bag, jumping between a variety of different issues.

Writing is good overall, some minor awkward phrasing etc.

Methodology & mastery of techniques

The student has clearly put in the work to get their head around the techniques and what they can do, and describes it well.

Results, data treatment & analysis

It takes awhile to clearly state how the techniques help address the problem of errors in single-cell-sequencing. But it becomes clear later on: basically they will look at measures of entropy in the posterior tree distribution, and try to identify taxa whose removal most contributes to reduction of entropy.

It is clear this could work as a way to rank sampled tips by their rogue-ishness. And the proposal discusses the challenge of e.g. rogue clades vs. rogue tips. Less discussed is the challenge of developing a rigorous “cut-off” for removing taxa. It might be that there is some clear separation between high-error and low-error single-cell sequences, but if not, the authors might have to invent something. I can imagine some rule-of-thumb methods from a (gasp!) frequentist perspective, a bit like methods for identifying outliers in linear regression analyses, but I imagine these might not be completely satisfying to a Bayesian lab group. Anyway, this proposal does not have to solve such issues, it will be interesting to explore the distribution of this entropy measure regardless.   
  
(I have to add: it may be there is some not-too-elaborate Bayesian approach? E.g. rather than trying to model the full error process in single-cell sequencing, perhaps each tip in the Bayesian analysis could be assigned a “weight” based on its observed entropy measure, and then these weights could be modified by some function with a parameter drawn from a prior & sampled during MCMC. Hopefully, high-entropy tips would be inferred to have low weight, and would then contribute less to the tree posterior. The “weight” would have to be something that effects the tip’s contribution to the likelihood – perhaps e.g.

* If the tip has an observed “C”, with the likelihoods at the tips being: 0 1 0 0
* A weight of 1 means the tips likelihoods are taken as-is.
* A weight of 0 means the likelihood at the tips is: 1 1 1 1, i.e. just like a “?” or “-“ code
* A weight of 0.5 would be something in-between, maybe just 0.5 1 0.5 0.5

…just a late-night brainstorm!)

Discussions & conclusions

Some good critical thinking, e.g. about the lack of known ground-truth, which is fine in this situation. The student should definitely think hard about these problems, there are a lot of potential avenues to explore to deal with unresolved problems like the above!

Organisation & presentation

Good.

Originality

Good.

Other comments (supervisor should comment here on degree of assistance required to complete experiments/produce the dissertation)

Name of examiner: \_Nick Matzke\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Signature: \_ \_\_

**POSTGRADUATE DISSERTATIONS AND PROJECTS**

**MARKING GUIDELINES**

|  |  |  |  |
| --- | --- | --- | --- |
| **Grade** | **Short comment** | **Explanation** | **Description** |
| A+ 90-100% | Exceptionally high level of performance | Fulfils all descriptors to an unusually high standard | * Well-structured and integrated research plan; * Well formulated research questions and appropriate investigative methodology used; * Excellent knowledge and understanding of subject; * Excellent knowledge of the literature and critical evaluation of previous work; * Significant new technical skills mastered * Appropriate data collected (if applicable) and rigorous and critical data analysis; * Clear understanding of significance of the data/evidence; * High level of critical thinking; * Arguments presented logically and coherently; * Conclusions comprehensive and well-justified; * Research project/dissertation well-constructed and appropriately illustrated. |
| A 85-89% | Clear high quality performance | Fulfils all descriptors to a high standard |
| A- | First class | Fulfils all descriptors to a high standard but not consistently |
| 80-84% |  |
| B+ | Very good | Fulfils all descriptors to a high standard by not consistently | * Sound research plan; * Good formulation of research questions and appropriate investigative methodology used; * Good grasp of subject matter; * Good knowledge of the literature and evaluation of previous work; * Appropriate data collected (if applicable) and good data analysis; * Some appreciation of the significance of the data/evidence; * Evidence of critical thinking; * Arguments presented reasonably well; * Some sound conclusions drawn; * Research project/dissertation reasonably well constructed and illustrated |
| 75-79% |  |
| B | Good | Fulfils most descriptors to a good standard |
| 70-74% |  |
| B- | Competent | Fulfils most descriptors to a good standard, but not consistently |
| 65-69% |  |
| C+ | Clear pass | Fulfils most descriptors to a competent standard | * Reasonable research plan outlined; * Research questions formulated and adequate research methodology applied; * Knowledge of subject matter shown but with some lapses, inadequacies and errors; * Adequate attempt at data analysis (if applicable) but may lack adequate justification; * Originality and critical thinking present but limited; * Arguments supported by some evidence; * Research project/dissertation adequately presented |
| 60-64% |  |
| C | Pass | Fulfils most descriptors to a competent standard but not consistently |
| 55-59% |  |
| C- | Marginal pass | Fulfils some descriptors to an adequate standard |
| 50-54% |  |
| D | Fail | Fails to fulfil enough descriptors to a competent standard | * Work lacks breadth and depth; |
| <50% |  | * Understanding and coverage inadequate; * No attempt at interpretation; |
|  |  | * Research project/dissertation poorly presented |