

# Third-year progress report for thesis advisory committee

This manuscript ([permalink](#)) was automatically generated from [mbhall88/TAC3\\_Report@c837c8d](#) on September 20, 2020.

## Authors

---

- **Michael B. Hall**

 [0000-0003-3683-6208](#) ·  [mbhall88](#) ·  [mbhall88](#)

EMBL-EBI; University of Cambridge · Funded by EMBL International PhD Programme (EIPP)

# Part A: Progress Report

---

## Examining bacterial variation with genome graphs

### Thesis Advisory Committee

- Zamin Iqbal (Supervisor) - EMBL-EBI
- John Marioni (Chair) - EMBL-EBI
- Georg Zeller - EMBL Heidelberg
- Estée Török - University of Cambridge

**Starting Date:** 12/10/2017

**Qualifying Assessment Date:** 06/07/2018

**Second TAC Meeting:** 15/10/2019

**Third TAC Meeting:** 13/10/2020

### Executive summary

Genomics is now ubiquitous in clinical and public health microbiology, at least in the developed world. However, many significant challenges remain. - Bacterial genomes harbour huge amounts of diversity, even within a species, and traditional reference-based approaches are problematic. - Much of the variation in bacteria is fundamentally inaccessible to short reads. - Long Nanopore reads are noisy, and SNP calling with this data is not properly benchmarked or standardised. - Since *Mycobacterium tuberculosis* (Mtb) infects so many people, there is potential for considerable impact for clinical applications. - There is also much to be gained from a high-quality pan-genome of Mtb as well as a detailed map of its enigmatic *pe/ppe* gene repertoire.

These motivations drive the following PhD thesis structure:

1. Develop algorithms and software for variant discovery using bacterial genome graphs, building on work of a previous student in the lab (my first paper, second author).
2. Benchmark Nanopore versus Illumina SNP calling, showing our algorithms meet the needs of clinical and public health users, validate, and publish (second paper).
3. Improve upon current whole-genome sequencing-based drug resistance prediction for Mtb using genome graphs.
4. Curate a high-quality reference pan-genome for Mtb that includes a detailed map of the *pe/ppe* genes.

## Part B: Training and career development

---

### Publication Strategy

Briefly outline the publication strategy; set priorities if needed.

### List of publications, papers in press, preprints, manuscripts submitted/in preparation to date

### Work plan and timeline for thesis submission

Provide a work plan to be completed prior to thesis writing and a timeline for submission.

### List of scientific courses and conferences attended to date and planned for next year

Please list all events in reverse chronological order.

### List of additional training, teaching and other relevant activities to date

Please list all training and other activities in reverse chronological order.

### Career development plan

**Please describe your current long-term career aims (i.e. 3-5 years after PhD).**

**Please comment on the types of position you would like to apply to for after the PhD and your expected application timeline?**

If applying for postdoc positions, which fields/fellowships are you considering? If you are applying for non-academic careers, what type of role? Do you have target companies or organisations in mind?

**What do you see as your strengths (2-3 skills)?**

**What do you see as your areas for improvement<sup>1</sup> (2-3 areas)?**

**What are your career development priorities until the end of your contract?**

Please take into account both the scientific and non-scientific skills you will need to work on a successful and timely completion of the PhD. List the skills you still need to acquire to achieve your longer-term career aims.

**What actions will you take to develop these skills?**

This may include training courses, opportunities to practice and develop these skills, or seeking feedback/guidance. N.B. fellows should take 1-2 non-scientific trainings or career development workshops per year, and at least one international conference during their PhD.

## Part C: Impact of COVID-19 pandemic

---

As outlined in the Procedures for COVID-19 related fellows' contract extensions, students and their GTL should document the impact of the corona crisis on the PhD project. The TAC should assess the approximate project delay and discuss how this could be mitigated over the remaining PhD period. Please use the space below to outline the impact of corona crisis on your project (e.g. lost productivity / project delays).

# References

---