**UNIVERSITY OF CAPE TOWN COMPUTATIONAL BIOLOGY DIVISION**

**Bioinformatics Support Request**

Please provide us with more information on your request for support. Complete the form as comprehensively as possible, and please indicate where there is still uncertainty.

**Please note, the earlier we are involved the better – for example, it would be better for us to be involved during the study design and even grant application stage**.

|  |  |
| --- | --- |
| **CONTACT DETAILS** | |
| Date of request | **23 January 2020** |
| Name | **Lucas Raphela** |
| Email address | [**RPHMAB004@myuct.ac.za**](mailto:RPHMAB004@myuct.ac.za) |
| Research Group/Department | **Molecular Mycobacteriology Research Unit** |
| Faculty | **Health Science** |
| IF student, name & email of supervisor | **Digby Warner**  [**Digby.warner@uct.ac.za**](mailto:Digby.warner@uct.ac.za) |

|  |
| --- |
| **PROJECT DETAILS** |
| 1. What is the scientific question? |
| Is there a genetic basis for the heterogenous colony morphologies observed upon repression of a riboflavin pathway gene in M. smegmatis? |
| 2. Who are the partners on the project? |
| None. |
| 3. What type of collaboration with CBIO is expected? For a project that is done as collaboration or for a fee, we will put the agreement in writing. |
| We would like to work collaboratively with a bioinformatician to optimize a pipeline for Mycobacterium smegmatis and Mycobacterium tuberculosis SNP calling to be adopted for future projects at the MMRU. |
| 4. Are there any ethical issues we should be aware of? |
| None. |
| 5. How much work is expected from CBIO and when? |
| We have conducted whole genome sequencing on two isolates with phenotypically different colony morphologies, in addition to sequencing the parental wild-type strain. We have the  sequencing files and would like a SNP analysis of the two isolates to be conducted against our reference wild-type genome. |
| 6. What type of data will be generated (e.g. sequencing, genotyping, expression, etc.) and what technology platform will be used? |
| We have WGS data generated using the Illumina platform. |
| 7. When do you expect the data? Does it need to be transferred from somewhere else? |
| The sequencing files are already available. Yes, the files are on a GenoHub server. |
| 8. How large will the data be? How long does it need to stored for, and have you made arrangements for storage? |
| I am not certain what size the files are. It can be stored on an external hard-drive. |
| 9. What bioinformatics analysis needs to be done? Which tools are required? |
| SNP calling. WGS pipeline with SAMtools or GATK variant caller. |
| 10. If a collaborative model is being used, what papers are envisaged and who will the authors be? |
| A single research paper with authors from the MMRU and contributors from CBIO. |
| 11. Can we add a short description and objective of the project to the CBIO website? |
| Yes. |

**PLEASE FORWARD THE COMPLETED FORM TO:**

[Nicola.mulder@uct.ac.za](mailto:Nicola.mulder@uct.ac.za)