**Project 2 Proposal: GitGud**

**Reviewers Comments**

Authors proposed using genomic mapping techniques to identify genomic variability in *Mycobacterium tuberculosis (M. tuberculosis),* thebacterial cause of tuberculosis*.* Significance of the project rests on the premise that currently used molecular tools in clinical settings are unable to correctly identify a special, recently appeared type of antibiotic resistance, referred to as cryptic resistance. Specifically, some strains of *M. tuberculosis* are falsely identified as antibiotic susceptible, while resistant to currently available antibiotics (false negative). Proposed genomic mapping should lead to characterization of structural variations (SVs) and single nucleotide polymorphism (SNPs) in such strains and shed some light on this phenomenon. The group wisely chose the mapping of the cryptically resistant strain to the reference genome of another, well characterized, complete, and virulent strain of *M. tuberculosis,* thus avoiding a possible confounding problem, had they used an avirulent reference strain. The proposal is well written, concise and focused.

The innovative part was the extension of the approach to utilize the SNPs to successfully attribute their contribution to resistance mechanisms. However, in the future writing, it is recommended to specifically pinpoint the ***significance*** and ***innovation*** ideas, as those are the key words the reviewers look for. Persistence of spelling of tuberculosis in first two paragraphs would give more professional impression (instead of lower case, upper case, abbreviation). The ***approach*** was in general very straight forward and clear with good reasoning why the pipeline components were selected. The only four suggested corrections would be 1) specifying if the reference genome (variant H37Rv) is antibiotic resistant or susceptible. The data interpretation will depend on this important detail that was not mentioned. 2), specific format of the data sets retrieved from NCBI should be specified. 3), the *fastQC* will need to be used again as a “b” in second steip in Figure 1, after the *Trimmomatic*. 4), Figure 1 was little bit difficult to read when printed black and white. Better contrasting of letters with respect to the background is recommended. ***Investigators*** should also specify a mechanism of resolving the disputes among themselves. Absolutely missing portion of the proposal was the **environment** to convince the reviewers the investigators have access to a high-performance computer cluster.