Whole genome sequencing of SARS-CoV-2: Adapting Illumina Protocols for quick and accurate outbreak investigation during A Pandemic

---

Abstract: The first COVID-19 detected case was in South Africa, then a few days after, an infection started in a large hospital outbreak in Durban, KwaZulu-Natal. Since then the pandemic has spread rapidly around the world. So we can use The Phylogenetic analysis of SARS-CoV-2 genomes to trace the path of transmission within a hospital. We also use it to identify strategies to know the source of the outbreak improve infection prevention and control. By doing this according to a plan we can genotype SARS-CoV-2 in near-real time during an urgent outbreak investigation. This plan included problems with the length of the original genotyping protocol, unavailability of reagents, and sample degradation and storage. So, three different library preparation methods for Illumina sequencing were set up to help completing the outbreak investigation in a few weeks. Which increased the success rate of sequencing the whole genome. A simple bioinformatics workflow for the assembly of high-quality genomes in near-real time was also fine-tuned.

Introduction: At the end of 2019, a mysterious pneumonia infected a number of people in Wuhan, China attributed to a new coronavirus. Despite the attempts to contain the virus in China, within a few months, the outbreak had reached and affected most countries around the world, including all countries in Africa. When SARS-CoV-2 takes place in a new graphical area, there must be important procedures which are: tracing and testing, detecting the causative agent and epidemiological investigations to control the infection outbreak and

identify clusters of patients. Genomic sequencing can be used identify the transmission routes of the virus in fast and accurate ways. This can then be used to trace the path of transmission within a population and to possibly identify the probable source, potentially leading to an improved public health response. This study outlines the processes involved in setting up genomic sequencing in order to investigate the COVID-19 outbreak . The process was more complicated than expected, as, in addition to data needing to be generated quickly, problems were also encountered with unavailability of stock, importing sequencing reagents, and sample degradation and storage. However, they managed to set up two new protocols which saved time to help the investigation.