

Concept Notes

Title: Molecular evolution and adaptation of SARS-CoV-2 omicron sub-lineage XBB Spike protein under African selection pressures.

Problem Statement: The SARS-CoV-2 omicron variant, first detected in South Africa in November 2021, has rapidly spread to many countries and regions, posing a serious threat to global public health. The omicron variant has more than 50 mutations in its spike protein, which may affect its transmissibility, immune evasion, and vaccine efficacy.

However, little is known about the evolution and adaptation of this variant in different geographic and epidemiological contexts. In particular, there is a lack of data and analysis on the mutations being acquired by the omicron sub-lineages circulating in Africa, where the variant originated and where there are diverse environmental and host factors that may exert different selection pressures on the virus.

Objective: This research project aims to identify and characterize the mutations being acquired by SARS-CoV-2 omicron sub-lineage XBB (XBB) in Africa due to the selection pressures specific to Africa. XBB is one of the most prevalent omicron sub-lineages in Africa, accounting for more than 40% of the omicron sequences from the continent as of January 2022. This research will provide insights into the evolutionary dynamics and adaptive potential of this sub-lineage and its implications for disease transmission, severity, immunity, and intervention strategies.

Approach: The approach of this research project will consist of three main steps:

- **Data collection:** We will collect genomic data of SARS-CoV-2 omicron sub-lineage XBB from various African countries and regions, using publicly available database GISAID. We will also collect relevant metadata such as sampling date, location, age, sex, vaccination status, clinical outcome, and co-infections.
- **Data analysis:** We will use bioinformatics tools and evolutionary biology methods to analyse the genomic data and identify the mutations being

acquired by XBB in Africa. We will use statistical models to infer the possible drivers and effects of the mutations, such as environmental factors, host immunity, viral fitness, antigenicity, and drug resistance.

- Data interpretation: We will interpret the results of the data analysis in light of the current scientific knowledge and literature on SARS-CoV-2 omicron variant. We will also discuss the limitations and uncertainties of our findings and suggest directions for future research. We will disseminate our results through peer-reviewed publications, conference presentations, and online platforms.

Work-plan: The expected duration of this research project is 1 month. The tentative work-plan is as follows:

- Week 1: Data collection and Data cleaning
- Week 2: Data analysis
- Week 3: Data interpretation
- Week 4: Preparing the manuscript for publication

Expected Outcomes: The expected outcomes of this research project are:

- A comprehensive dataset of SARS-CoV-2 omicron sub-lineage XBB genomic sequences and metadata from Africa
- A detailed characterization of the mutations being acquired by XBB in Africa due to the selection pressures specific in Africa.