Transmission analysis using 101 SNPs barcode data

27 March, 2024

## Data Source

This analysis aims to use *101* SNPs barcode data and *2498* isolates to explore Plasmodium *falciparum* drug resistance profile in the Gambia.

## Analysis steps

1. Genetic diversity analysis
   * Allelic richness
   * Mutations / Haplotypes Allele frequencies
2. Transmission dynamics analysis
   * Identity by state
   * Jaccard Similarity Coefficient
   * Sørensen-Dice Coefficient (Dice Similarity)
   * FST (Fixation Index)
   * Nei’s Genetic Distance (Nei 1972)
3. Temporal dynamics analysis
   * Temporal FST (Fixation Index)
   * Haplotype Diversity Over Time
   * Drug Resistance Analysis

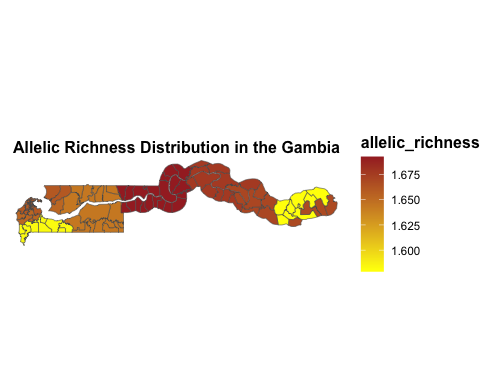
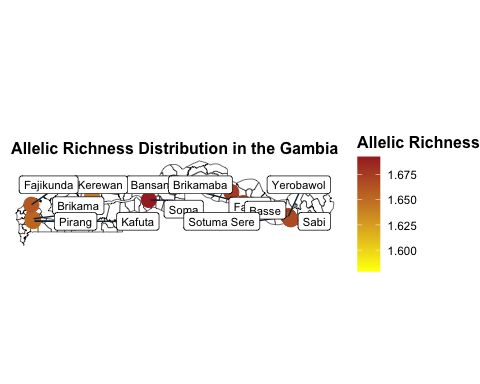
### Genetic diversity analysis

Assess the genetic diversity of Plasmodium *falciparum* across different villages and time points. This includes calculating measures such as allelic richness, and SNP frequencies.

#### Allelic richness

Allelic richness is a measure of the number of alleles per locus in a given population, adjusted for the smallest sample size if there are varying sample sizes.

FALSE Reading layer `geoBoundaries-GMB-ADM3\_simplified' from data source   
FALSE `/Users/mdiop/Library/CloudStorage/OneDrive-LondonSchoolofHygieneandTropicalMedicine/PhD/Analysis/01\_PhD\_project/01\_analysis/01\_objective1/barcodes/gis\_coordinates/geoBoundaries-GMB-ADM3-all/geoBoundaries-GMB-ADM3\_simplified.shp'   
FALSE using driver `ESRI Shapefile'  
FALSE Simple feature collection with 120 features and 5 fields  
FALSE Geometry type: POLYGON  
FALSE Dimension: XY  
FALSE Bounding box: xmin: -16.8142 ymin: 13.04347 xmax: -13.79112 ymax: 13.82324  
FALSE Geodetic CRS: WGS 84



### Transmission dynamics analysis

Identify genetic clusters of P. *falciparum* using methods like network analysis to infer transmission links. Create Network from genetic similarity indices (IBS, Jaccard Coefficient, Dice Similarity, Fst and Nei’s Genetic Distance)

#### Identity by state

#### Jaccard Similarity Coefficient

#### Sørensen-Dice Coefficient (Dice Similarity)

#### FST (Fixation Index)

#### Nei’s Genetic Distance (Nei 1972)

### Hotspot Identification

Use spatial analysis to identify areas with high transmission rates or clusters of genetically similar infections. Spatial analysis to identify hotspots of high transmission rates or clusters of genetically similar infections involves using statistical methods and indices that can detect areas with significantly higher incidences of disease or genetic similarity than would be expected by chance. This analysis can be crucial for targeting interventions and understanding the spatial dynamics of infectious diseases.

### Temporal dynamics analysis

To analyze how genetic variation changes over time and understand the spread and dynamics of transmission, several indices and methods can be employed. These analyses help in tracking the evolution of pathogens, identifying the emergence of new variants, and understanding how these changes influence transmission dynamics.

#### Temporal FST (Fixation Index)

Temporal FST compares genetic differentiation between populations at different time points. It measures how genetic variance is distributed over time, providing insights into the population structure’s temporal dynamics. A high FST value indicates significant genetic differentiation, which could result from selection pressure, population bottlenecks, or founder effects.

#### Haplotype Diversity Over Time

Haplotype diversity measures the uniqueness of the genetic variants present in a population. Tracking changes in haplotype diversity over time can indicate the emergence or disappearance of strains, reflecting how genetic diversity is influenced by transmission dynamics, selection pressures, and population size changes.

### Drug Resistance Analysis

Since the SNPs are linked to drug resistance, we analyzed the prevalence of these resistance-associated SNPs in different locations and over time.

Estimating the prevalence of molecular markers associated with drug resistance involves identifying specific genetic mutations or patterns within a pathogen’s genome that confer resistance to antimicrobial or antiparasitic drugs. This process is crucial in understanding the spread of resistance and informing treatment guidelines.

