**Dengue virus**

Dengue is a virus of the flaviviral family (1,2), sometimes referred to as arbovirus, which stands for arthropod-borne transmitted virus. These vectors are mosquitoes of the *Aedes* genus and two species *A. aegypti* and *A. albopictus* are the ones responsible of transmission. Both species have a specialized ecological niche, *aegypti* is responsible for urban transmission and likely specializes in feeding on human blood and breeding in human-made water bodies approximately 5000 years ago in Africa, and *albopictus* feeds mainly on animals (including humans), and breeds on natural water bodies thus is more commonly found in the sylvatic transmission cycle (3).

These 4 viruses cause mild or no symptoms on the first infection but on the secondary infection disease is severe by a mechanism named Antibody Dependent Enhacement (ADE) (4). The severity of disease on the secondary infection depends on the amount of DENV antibodies built on the first infection and with a narrow range of DENVab resulting in severe disease (4). However, secondary infections with viruses of the same serotype as the one of the first infection are impossible, and this is named homotypic (same serotype) lifelong cross-protection. The heterotypic secondary infections (with a different serotype from the one of the first infection) are thus the ones that result in disease.

The antigenic map (Figure 1) is a 2D visualization of an assay that quantifies how much protection a secondary infection confers on a first-infecting virus. The position in the map is a phenotype for that respective strain, an antigenic phenotype. In the homework, we want to ask if genetic data have information about this phenotype. Because the genotype to phenotype problem (sequence to position in the map, Figure 1) is extremely hard, we will rather ask a slightly different question: does the genetic difference explain the difference in this antigenic phenotype?

In the first part of the homework (point 2), I want to use a distance between a pair of sequences to predict the distance in the antigenic map. As an example, in Figure 1, I draw a black dashed line to show the distance between those two pairs of strains. The distance between a pair of strains I want you to use as an explanatory variable is the Percentage Identity Matrix (PID). The PID measures the nucleotide divergence between two sequences: how many bases the sequences have in common divided by the sequence length?

Intragenotypic variability (see point 3 in the homework) exists, and thus, in the second part, I want you to *control* for some of this variability using as a second predictor the distance of each strain to the center of its respective phenotype. For the third part, I built a tree, and I want you to use the time to a most recent common ancestor (TMRCA) between a pair of strains as the additional descriptor for the 3rd linear model.

The linear models and its equations are shown below.

A graph of different colored dots

Description automatically generated

Figure 1. **Antigenic phenotype.** Each dot represents a virus/strain; DENV serotypes are color-coded as indicated in the legend. Two strains from serotype 4 (purple) and serotype 1 (red) are highlighted, as well as the antigenic distance between that pair of strains.

A diagram of different colored dots

Description automatically generated

Figure 2. **Center of each serotype.** Strains are presented in Figure 1; center of each serotype is highlighted with a gold star as indicated in the legend.

**References**

1. Mifsud JCO, Lytras S, Oliver MR, Toon K, Costa VA, Holmes EC, et al. Mapping glycoprotein structure reveals Flaviviridae evolutionary history. Nature [Internet]. 2024 Sep 4 [cited 2024 Sep 4]; Available from: https://www.nature.com/articles/s41586-024-07899-8

2. Holmes E, Twiddy S. The origin, emergence and evolutionary genetics of dengue virus. Infection, Genetics and Evolution. 2003 May;3(1):19–28.

3. Rose NH, Badolo A, Sylla M, Akorli J, Otoo S, Gloria-Soria A, et al. Dating the origin and spread of specialization on human hosts in Aedes aegypti mosquitoes. eLife. 2023 Mar 10;12:e83524.

4. Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, et al. Antibody-dependent enhancement of severe dengue disease in humans. Science. 2017 Nov 17;358(6365):929–32.