Detection and recontruction of viral haplotypes from APMV-1 samples

Giovanni Marques de Castro[1], Francisco Pereira Lobo[1], Helena Lage Ferreira[2]

1 UFMG
2 USP

Abstract

Two samples of APMV-1 were sequenced using MiSeq and analyzed to verify which genotype the samples belong. As viruses have very high error rates when replicating it is expected to find variants given a high depth of sequencing as provided by NGS. The data generated for both samples were enough to analyze the underlying viral population, the quasispecies of APMV-1 for both samples. For this, the reads for each library were mapped to a few reference genomes and selected that which had the most mapped reads, the KJ123642. Using the aligned reads as the input for the software QuasiRecomb and restricting the region of to the F protein, the haplotypes for both samples were reconstructed. The most frequent haplotype from each sample and other 88 APMV-1 genomes from NCBI were used to reconstruct the phylogeny, the analysis of the phylogeny allow to visualize in which genotype they belong. The F protein is know to have a cleavage site in which the amino acid present can be used to predict if the virus is lentogenic or velogenic. The analysis of the cleavage site revealed that the most frequent haplotypes from both samples are velogenic. More than 65% of the reads were aligned to the reference genome (KJ123642) for each sample. The phylogenetic analysis showed that they group with the Vb genotype. Using more of the reconstructed haplotypes to reconstruct the phylogeny showed an extremely close result to using only the most abundant haplotype, clustering together, most likely due to the founder effect of a small related viral population.

Funding: Capes