Evolutionary genomics of SARS-CoV-2: keys for viral success

Background: Deciphering the impact of SARS-CoV-2 mutations in immune evasion is essential to understand the dynamics of the epidemic. We aim to better understand the potential of different SARS-CoV-2 variants for impacting transmissibility and virulence.

Methods: We employ a complex workflow starting with mutation detection and epidemiological investigation, to further characterise the impact of mutations in the structure of the SARS-CoV-2 proteins and in the biology of a pseudovirus system in vitro (Figure 1). Mutation surveillance has employed different datasets including a global dataset with more than seven million viral sequences, and three local datasets up to ten thousand sequences paired with detailed hospitalisation and vaccination information.

Results: First we have identified cluster 1163.7, characterised by spike mutations S:D1163Y and S:G1167V inside cluster 20E and we concluded that the two spike mutations impacted syncytia formation and infectivity in vitro. However, they did not impact thermal stability or neutralisation by antibodies except from those infected in the first wave. Second, we have studied which combinations of S1 terminal domain mutations present in two variants of concern are more transmissible in different epidemiological settings. Third, we have studied the diversity and migration dynamics of variants of concern. Finally, we have used different statistical approaches to find associations between viral genomic variants and vaccine breakthrough and/or hospitalisation.

Conclusions: We have characterised SARS-CoV-2 variants circulating globally, with special focus in Spain, and studied the association of specific viral variants with vaccine breakthrough, hospitalisation, and transmission. Our study provides an effective pipeline for the characterization of biological fitness driven by the identification of mutations of relevance