

Integrated genomic and structural analysis of SARS-COV-2, angiotensin converting enzyme 2 receptor (ACE2) and transmembrane protease serine 2 (TMPRSS2) in SARS-COV-2 tested patients at a tertiary hospital in Gauteng

¹Seakamela ME, ^{1,2}Moeng K, ^{1,2}Selabe SG, ¹Gededzha MP

¹Department of Virology, Sefako Makgatho Health Sciences University, Pretoria, South Africa; ²National Health Laboratory Service, Department of Virology, Sefako Makgatho Health Sciences University, Pretoria, South Africa

Background

SARS-CoV-2 infection outcomes are influenced by viral and host genetic factors, particularly genes involved in viral entry such as angiotensin converting enzyme 2 receptor (ACE2) and transmembrane protease serine 2 (TMPRSS2). However, data is limited on the integrated analysis of viral and host genetic variation in African populations. This study aimed to characterize SARS-CoV-2, ACE2 and TMPRSS2 in nasopharyngeal samples of patients tested for SARS-CoV-2 at a tertiary hospital in Gauteng.

Methods

This study analysed 110 residual nasopharyngeal swab samples collected from patients between 2020 and 2024. SARS-CoV-2 spike, ACE2 and TMPRSS2 genes were amplified and sequenced using Next generation and sanger sequencing. Phylogenetic analysis was performed using Nextclade and MEGA. Functional implications of mutations were assessed using SIFT, glycosylation and phosphorylation prediction, molecular docking to evaluate protein-protein interactions. Statistical analysis was performed with STATA.

Results

SARS-CoV-2 spike analysis revealed co-circulation of multiple lineages, with B.1.2 predominance. The D614G mutation was the most frequent and was present across all clades. Several mutations, including H655Y, N679K, and P681H, were identified although most substitutions were predicted to be tolerated.

ACE2 were highly conserved, while TMPRSS2 exhibited greater genetic variability. A rare ACE2 variant (T34664C) was detected in SARS-CoV-2-positive patients and was predicted to affect protein function. TMPRSS2 variants, (G27589A and G27774T), showed statistical association with infection status and were predicted to have potential functional effects.

Post-translational modification analysis revealed largely conserved glycosylation and phosphorylation patterns across all proteins. Molecular docking demonstrated strong binding between SARS-CoV-2 spike and ACE2 and TMPRSS2 and ACE2 variants.

Conclusion

This study demonstrated viral and host genetic variability in a South African population with ACE2 and TMPRSS2 variants contributing to viral entry dynamics. Integrated viral-host genomic and structural analyses provide important insights into SARS-CoV-2 pathogenesis, although experimental validation is required to confirm predicted functional effects.