Variations in the S Segment of Rift Valley Fever Virus with Special Reference to the Nonstructural NSs Coding Region

Susan Claire Aitken

Rift Valley fever virus (RVFV) is a *Phlebovirus* member of the *Bunyaviridae* family and it is the causative agent of Rift Valley fever (RVF), a mosquito-borne viral zoonotic disease that poses a significant threat to domestic ruminants and human health in Africa. The RVFV is an encapsulated, negative-sense, single-stranded RNA virus with a tripartite segmented genome, containing L (large), M (medium) and S (small) segments. The S segment codes for two proteins, namely the nucleocapsid (N) protein and non-structural protein (NSs). There is evidence that the NSs protein is involved in virulence by blocking the expression of the interferon beta (IFN-β) promoter. It has been recently demonstrated that the SAP30-NSs-YY1 multiprotein complex represses the IFN-β promoter. Consequently, the interferon expression is blocked, allowing virus to replicate.

A total of 45 isolates of RVFV recovered over a period of 53 years in 14 African countries, Madagascar and Saudi Arabia were characterized by full sequencing of the S segment of the virus. This data was added to another 27 strains of RVFV available on GenBank for phylogenetic analysis using MEGA4, giving a total of 72 strains analyzed. Alignments were made of the entire S segment, the NSs gene, the N gene, and their deduced amino acid sequences. The laboratory strains, clone 13, MP12 and Smithburn, were also included in the alignments.

Two isolates were passaged ten times through two different amplification systems to asses the potential for sequence variation to occur in the original material through routine laboratory manipulations. Sequencing data was generated from the virus RNA present in the original clinical specimens and from the extracted RNA from the tenth passage of virus in each amplification system. The results showed 100% homology for each respective isolate, demonstrating that the RVFV S segment remained stable during ten serial passages in different propagation systems.

Phylogenetic analysis was conducted on the naturally occurring RVFV strains (n = 72) and the findings indicate that circulating strains are compartmentalized and belong to one of three major lineages, namely Egyptian, western African, and central, eastern and southern African. The strains clustered in the Egyptian lineage had an average p-distance of 1.0%, the western African strains 0.9%, and the central, southern and eastern African strains 2.0%. The overall average p-distance was 2.5%, with a range from 0 to 4.1%. For the N gene, the range was from 0 to 4.2%, with an average of 2.2%. For the N protein, the range was from 0 to 2%, with an average of 0.2%. The NSs gene had a range of 0 to 4.6%, with an average of 2.4%. The NSs protein had a range of 0 to 3.8%, with an average of 1.7%. The intergenic region (IGR) had a range of 0 to 9.2%, with an average of 4.8%.

Results of the study suggest that RVF outbreaks can result from either the rapid spread of a single strain over vast distances or from an increased activity of a strain circulating at an endemic level within an area/region during prolonged dry periods.

Sequencing alignment showed that the length of the S segment ranged from 1690 to 1692 nucleotides. This difference in length was due to insertions and deletions found in the IGR, which is also the region with the most sequence divergence (4.8%). Both the NSs and N genes had neither insertions nor deletions, and were both found to be stable, though the NSs gene was slightly more variable than the N gene (2.5% versus 2.2%)

The deduced amino acid sequences of the NSs protein were considerably more variable than that of the N protein (1.7% versus 0.2%). Alignment of the NSs protein demonstrated that the 5 cysteine residues at positions 39, 40, 150, 179 and 195, are highly conserved among the isolates analyzed. These residues are important for conservation of the three-dimensional structure of the protein and the formation of filamentous structures observed in cells infected with natural strains of RVFV. The NSs protein is now implicated as the major factor of virulence and that its pathogenicity is associated with the blocking of interferon production. Therefore, any amino acid changes that result in changes to the filamentous structure of the NSs protein might impact on the binding kinetics between the NSs protein, SAP30 (Sin3A

Associated Protein 30) and YY1 (Yin Yang-1). There were 6 amino acid changes in the NSs-SAP30 binding domain, with one being unique to the live-attenuated Smithburn vaccine strain.

Generated sequencing data contributes to global phylogenetic characterization of RVFV isolates and molecular epidemiology of the virus. In addition, findings of this study will further aid investigation on reassortment events occurring between strains of RVFV and genetically related viruses, the role of the NSs protein in the replicative cycle of the virus, the pathogenic effects of the NSs protein within the RVFV-infected host cells, and might help to identify molecular basis of RVFV virulence.