Characterization of the Salehabad virus species complex of the genus Phlebovirus (Bunyaviridae)

Gustavo Palacios,1,2 Nazir Savji,2† Amelia Travassos da Rosa,3 Aaloki Desai,2 Maria Paz Sanchez-Seco,4 Hilda Guzman,3 W. Ian Lipkin2 and Robert Tesh3

1Center for Genomic Sciences, United States Army Medical Research Institute for Infectious Diseases, Frederick, Maryland, USA
2Center for Infection and Immunity, Columbia University, New York, New York, USA
3Center for Biodefense and Emerging Infectious Diseases, Department of Pathology, University of Texas Medical Branch, Galveston, TX, USA
4National Center for Microbiology, Instituto de Salud ‘Carlos III’, Madrid, Spain

Genomic and antigenic characterization of the Salehabad virus, a species of the genus Phlebovirus, and four other unclassified phleboviruses (Arbia, Adria, Arumowot and Odrenisrou) demonstrate a serological and genetic relation to one another and are distinct from the eight other recognized species within the genus Phlebovirus. We propose to incorporate these four unclassified viruses as part of the Salehabad species complex within the genus. The known geographical distribution for the members of this species group includes southern Europe, Central Asia and Africa.

The family Bunyaviridae comprises five genera: Hantavirus, Nairovirus, Orthobunyavirus, Phlebovirus and Tospovirus that are differentiated by antigenic and molecular characteristics (Nichol et al., 2005). Viruses in the genus Phlebovirus are further subdivided into two groups: the Sandfly fever and the Uukuniemi groups, based on their antigenic relationships, arthropod vectors and the presence or absence of a non-structural protein in the M segment (Nichol et al., 2005). Phlebotomine sandflies and mosquitoes transmit viruses in the Sandfly fever group, while ticks transmit Uukuniemi group viruses. A number of the phleboviruses are associated with human illness and at least one, Rift Valley fever virus, is a major veterinary pathogen (Peters et al., 2011; Tesh, 1989; Yu et al., 2011).

Because of the paucity of genetic data for most of the named phleboviruses, the International Committee on Taxonomy of Viruses defines species within the genus Phlebovirus by their antigenic relationships (Nichol et al., 2005). Cross-complement fixation (CF) and plaque reduction neutralization tests have generally been used for the classification of viruses in this genus (Tesh et al., 1976, 1982; Travassos da Rosa et al., 1983). The present communication is the fourth report in our effort to develop a more precise classification system for the phleboviruses by sequencing available viruses in the genus and reassessing their antigenic relatedness in order to clarify their phylogenetic relationships (Palacios et al., 2011a, b; 2012). Here, we describe the genetic and antigenic relatedness of another of those recognized species groups, Salehabad, that contains Salehabad (SALV), Arbia (ARBV), Adria (ADRV), Arumowot (AMTV) and Odrenisrou (ODRV) viruses.

The phlebovirus strains used in this study were AMTV strain Ar 1286-64; SALV, strain I-81; ARBV, strain ISS. Phl.18; and ODRV strain ArA 1131/80. ADRV was not available for study, as noted before. Stocks of AMTV, SALV, ARBV and ODRV were obtained from the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) at the University of Texas Medical Branch, USA.

AMTV was originally isolated in 1963 from Culex antennatus mosquitoes collected in Sudan. Subsequently, it has been isolated from mosquitoes and rodents throughout Africa. Although antibodies to AMTV have been reported in serosurveys in African human populations (Tesh et al., 1976), there is no definitive evidence that it causes human illness.

SALV was isolated from a pool of sandflies (Phlebotomus spp.) collected in a rural village in central Iran in 1959. No subsequent recoveries of this virus have been made and it is unknown whether it is associated with illness in humans or domestic animals.

†Present address: School of Medicine, New York University, New York, New York, USA.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are JX472400–JX472405.

Supplementary figures are available with the online version of this paper.

Received 9 October 2012
Accepted 6 December 2012

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are JX472400–JX472405.

Supplementary figures are available with the online version of this paper.
ARBV was originally isolated from a pool of *Phlebotomus perniciosus* collected near Toscana, Italy in 1980 (Verani *et al.*, 1988). Subsequent recoveries of the virus have been made in the same region from other sandfly species, but ARBV has not yet been associated with human illness.

ADRV has not yet been isolated. Partial RNA of a novel phlebovirus, related to ARBV was first detected in two pools of phlebotomine sandflies (species unknown) collected on the Adriatic coast of Albania in 2005 (Papa *et al.*, 2011). The same RNA sequence was subsequently detected in the blood of a child hospitalized with a febrile illness and seizure in northern Greece in 2009 (Anagnostou *et al.*, 2011).

ODRV was isolated from a single pool of *Culex albiventris* mosquitoes collected in the Tai forest, Ivory Coast in 1980 (CDC, 2013; http://wwwn.cdc.gov/arbocat/catalog-listing.asp?VirusID=342&S1=1). It has not yet been associated with human disease.

Methods used to prepare antigens and antibodies (immune ascitic fluids) have been described previously (Beaty *et al.*, 1989; Travassos da Rosa *et al.*, 1983; Huang *et al.*, 2007). All immunizing antigens and antibodies were prepared in mice.

Antigens for serological tests were prepared from infected newborn mouse brain (AMTV and ODRV), or spent medium from infected Vero cell cultures (SALV and ARBV) was used. CF tests were performed by the microtitre technique (Beaty *et al.*, 1989; Xu *et al.*, 2007). HI tests were performed with 4 U haemagglutination of virus at the optimal pH (5.75) against serial twofold antiserum dilutions starting at 1 : 20. HI titres were recorded as the highest dilutions giving 3 + or 4 + fixation of complement. Haemagglutination-inhibition (HI) tests were done in microtitre plates, as described previously (Travassos da Rosa *et al.*, 1983). HI tests were performed with 4 U haemagglutination of virus at the optimal pH (5.75) against serial twofold antiserum dilutions starting at 1 : 20. HI titres of 1 : 20 or greater were considered positive. All serological tests were performed at the University of Texas Medical Branch, Galveston, USA.

Results obtained in HI and CF tests with SALV, ARBV, AMTV and ODRV are summarized in Table 1. Since ADRV has not yet been isolated and is known only from a partial sequence of its large (L) RNA segment, it was not included in the serological tests. The antigens of SALV, ARBV, AMTV and ODRV were reactive with their homologous antisera. In CF tests, SALV and ARBV antigens and antisera were highly cross-active, as were AMTV and ODRV; in fact, the four viruses formed two distinct serological groups by this method. In HI tests, the two serological groupings were again observed, although ARBV antigen reacted with AMTV and ODRV antibody and ARBV antigen reacted with AMTV and ODRV antigens. On the basis of these cross-reactions, we feel that the four viruses constitute a single (Salehabad) antigenic complex.

For genome sequencing, viral stocks were extracted using TRIzol LS (Invitrogen). Total RNA extracts were treated with DNase I (DNA-Free; Ambion). cDNA was generated using the Superscript II system (Invitrogen) employing random hexamers linked to an arbitrary 17-mer primer sequence (Palacios *et al.*, 2007). Resulting cDNA was treated with RNase H and then randomly amplified by PCR with a 9 : 1 mixture of primer corresponding to the 17-mer sequence and the random hexamer linked 17-mer primer (Palacios *et al.*, 2007). Products greater than 70 bp were selected by column chromatography (MiniElute; Qiagen) and ligated to specific adapters for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences) without fragmentation (Cox-Foster *et al.*, 2007; Margulies *et al.*, 2005; Palacios *et al.*, 2008). Software programs accessible through the analysis applications at the GreenePortal website (https://tako.cpmc.columbia.edu/Tools/) were used for the removal of primer sequences, redundancy filtering and sequence assembly. Sequence gaps were completed by PCR, using primers based on pyrosequencing data. Amplification products were size-fractionated on 1 % agarose gels, purified (MiniElute; Qiagen) and ligated to specific adapters for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences) without fragmentation (Cox-Foster *et al.*, 2007; Margulies *et al.*, 2005; Palacios *et al.*, 2008). Software programs accessible through the analysis applications at the GreenePortal website (https://tako.cpmc.columbia.edu/Tools/) were used for the removal of primer sequences, redundancy filtering and sequence assembly. Sequence gaps were completed by PCR, using primers based on pyrosequencing data. Amplification products were size-fractionated on 1 % agarose gels, purified (MiniElute; Qiagen) and directly sequenced in both directions with ABI Prism Big Dye Terminator 1.1 Cycle Sequencing kits on ABI Prism 3700 DNA Analysers (Perkin-Elmer Applied Biosystems). For the termi of each segment, a primer with the 8 nt conserved sequence was used for a specific reverse transcription with

| Table 1. Salehabad antigenic complex: results of CF and HI tests |

<table>
<thead>
<tr>
<th>Antibody</th>
<th>CF test</th>
<th>HI test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SALV</td>
<td>ARBV</td>
</tr>
<tr>
<td>Salehabad</td>
<td>256/ ≥Φ</td>
<td>≥256/ ≥Φ</td>
</tr>
<tr>
<td>Arbia</td>
<td>128/ ≥Φ</td>
<td>512/ ≥Φ</td>
</tr>
<tr>
<td>Arumowot</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Odrenisrou</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*CF titres are expressed as highest antibody/highest antigen dilution. 0 = <8/Φ.
*ΦUndiluted.
+HI titres are expressed as highest antibody dilution. 0 = <1 : 20.*
additional arbitrary nucleotide on the 5' end (5'-AAGCAG-TGGTATCAACGCAGTACACACAAAG-3'; the bold portion is the conserved nucleotides). This primer is designed to bind to the 3' end of the genomic RNA and the 3' end of the mRNA. Sequences of the genomes were verified by classical dideoxy sequencing using primers designed to create products of 1000 bp with 500 bp overlap from the draft sequence.

SALV, ARBV, AMTV and ODRV all have the genomic organization characteristic of phleboviruses: three RNA segments including a L segment encoding the RNA polymerase, a medium (M) segment encoding a polyprotein that includes the non-structural protein (NSm) and both glycoproteins (GN and GC), and a small (S) segment encoding the N protein (NP) and, in ambisense orientation, a non-structural protein (NSs) (GenBank accession nos HM566143–HM566145, HM566173–HM566173, JX472400–JX472405).

All phlebovirus sequences available from GenBank (March 2012) was used to determine the phylogenetic relationships of SALV, ARBV, AMTV and ODRV. All sequences were aligned using the CLUSTAL algorithm (as implemented in the MEGA package version 3) at the amino acid level with additional manual editing to ensure the highest possible quality of alignment. Neighbour-joining analysis at the amino acid level was performed due to the observed high variability of the underlying nucleotide sequences. Nucleotide trees were achieved by using neighbour-joining analysis and the Kimura 2-paramater model. The statistical significance of tree topology was evaluated by bootstrap resampling of the sequences 1000 times. Phylogenetic analysis was performed by using the MEGA software (Kumar et al., 2004).

In phylogenetic analysis, major nodes that represented viruses belonging to the same species or antigenic complex were clearly distinct and confirmed previously reported topologies (Charrel et al., 2009; Collao et al., 2009). Although the four viruses reported here clustered together defining the Salehabad virus antigenic complex node (Figs 1 and 2, full trees in Figs S1, S2 and S3, available in JGV Online), only the NS ORF showed a strong bootstrapping value. Interestingly, the serological pairings (SALV and ARBV; AMTV and ODRV) were strongly supported by the phylogenetic analysis. Similar topologies were observed in phylogenetic trees at the nucleotide level (data not shown). No branching inconsistencies were observed suggesting no reassortment among these viruses.

For sequence assembly and analysis, Geneious 4.8.3 (Biomatters Inc.) was used. Topology and targeting predictions were generated by employing SignalP, NetNGlyc, TMHMM (http://www.cbs.dtu.dk/services), TopPred2 (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html), and integrated predictions in Geneious (Bendtsen et al., 2004; Claros & von Heijne, 1994; Kahsay et al., 2005; Käll et al., 2004; Krogh et al., 2001)

ORF analysis of the RNA-dependent RNA polymerase (L) revealed that the four members of the Salehabad virus antigenic complex show high conservation with previously described conserved phlebovirus functional domains (Aquino et al., 2003; Palacios et al., 2011b; Poch et al., 1989; Xiong & Eickbush, 1990). The polyprotein is co-translationally cleaved into three protein products, NSm, Gn and Gc. The topology of the polyprotein of the Salehabad virus antigenic complex is similar to the predicted topology for TOSV, PTV, RVFV and CDUV (Gerrard & Nichol, 2002; Grø et al., 1997; Ihara et al., 1985; Matsuoka et al., 1996; Palacios et al., 2011b; Valentini et al., 2008) with signal sequences, transmembrane domains and predicted cleavage sites conserved (data not shown). Similarly, all the functional domains described in the NP (Gerrard & Nichol, 2002; Le May et al., 2005; Palacios et al., 2011b) were conserved (data not shown). No singularities were found in the NSs ORF.

With the advent of high-throughput sequencing, virus taxonomy has begun to rely less on serology and more on genetic characteristics. One of the problems with antigenic or serological classification is the variability in the reactivity and potency of each batch of antigen or antibody. A second problem, in the case of the phleboviruses, is that some of the viruses are not pathogenic in mice, so production of ‘clean’ murine antiserum is difficult. Also as more and more investigators use RT-PCR and other RNA detection systems for virus assay, virus isolates are sometimes not available for serological study, as in the case of ADRV. Genetic classification provides a more objective standard from which to classify viruses, with numerical cut-offs to distinguish different taxonomic levels. Furthermore, genome information may be the only information available, as in the case of ADRV.

However, a genetics-only based classification system for the bunyaviruses also has risks. Among phleboviruses, we have demonstrated that serological relatedness can sometimes be detected in cases where no genetic similarity is observed (Palacios et al., 2012). The serological and genetic data among the members of the Salehabad group enhance that view. Taken as a whole, we feel there is evidence to group AMTV, SALV, ARBV, ADRV and ODRV, as a single species group or complex in the genus Phlebovirus that is geographically distributed in Africa, Asia and Europe. Sequence or serological data alone show only a slight association; together the two datasets support this clustering. This data correlates and emphasizes our recent observation on the Uukuniemi antigenic complex, suggesting that a system of classification for the family Bunyaviridae that considers both genetic and serological data, when available, is optimal.

Salehabad has been detected in the Mediterranean region in clinically ill persons and sandflies. In the same geographical region, human phlebovirus seroprevalence is high. Moreover, a Sicilian-like virus (Cyprus virus) was associated with a major outbreak of febrile illness in 2002.
Characterization of the genetic diversity of the members of this group might be crucial to enhance molecular diagnostic methods in the area.

Acknowledgements

This work was supported by Google.org, National Institutes of Health award AI57158 (North-east Biodefence Center - Lipkin), and USAID Predict funding source code 07-301-7119-52258 (Center for Infection and Immunity) and the Department of Defense. A. T. R., R. T. and H. G. were supported by NIH contract HHSN2722010000ROI/HHSN27200004/DO4.

References


http://vir.sgmjournals.org

Fig. 1. Phylogenetic analysis of select phlebovirus (a) L, (b) M, (c) N and NS ORFs. Neighbour-joining analysis at the amino acid level was performed due to the high variability of the underlying nucleotide sequences. The statistical significance of the tree topology was evaluated by bootstrap resampling of the sequences 1000 times.

Fig. 2. Phylogenetic analysis of the Adria virus L sequence. Neighbour-joining analysis at the amino acid level was performed due to the high variability of the underlying nucleotide sequences. The statistical significance of the tree topology was evaluated by bootstrap resampling of the sequences 1000 times.


