The S RNA genome segments of Batai, Cache Valley, Guaroa, Kairi, Lumbo, Main Drain and Northway bunyaviruses: sequence determination and analysis

Ewan F. Dunn, David C. Pritlove† and Richard M. Elliott*

Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, U.K.

Bunyaviruses have a genome comprising three segments of negative-sense RNA. The smallest RNA segment, S, encodes the nucleocapsid protein, N, and a non-structural protein, NSs, in overlapping reading frames. The sequences of the S genome RNA segments of seven bunyaviruses (Batai, Cache Valley, Guaroa, Kairi, Main Drain, Northway and Lumbo) were determined from cloned cDNAs obtained using a one-step reverse transcription-PCR protocol. These sequences were compared to those of six viruses previously published, reinforcing earlier conclusions about relationships of the bunyaviruses. Sequence homologies between N proteins correlated with the subdivision of these viruses into three serogroups, Bunyamwera, California and Simbu.

Introduction

Viruses in the family Bunyaviridae are characterized by the possession of a genome comprising three segments of ssRNA designated L (large), M (medium) and S (small). The 300 or so viruses in the family are classified into five genera: Bunyavirus, Nairovirus, Hantavirus, Phlebovirus and Tospovirus. The Bunyavirus genus is the largest of the genera in the Bunyaviridae and contains at least 162 named viruses which are subdivided into 18 serogroups on the basis of complement fixation (CF), haemagglutination-inhibition (HI) and neutralization (NT) antibody studies (Calisher, 1991). Few bunyaviruses (i.e. members of the Bunyavirus genus) have been characterized at the molecular level, but it has been determined that the L RNA encodes the L protein (RNA polymerase); the M segment codes for a polyprotein precursor of the two virion glycoproteins, G1 and G2, and a non-structural protein NSm; and the S RNA segment encodes two proteins, N and NSs, in overlapping reading frames (ORFs), which are translated from the same complementary mRNA as the result of alternative initiation of translation (for reviews, see Bishop, 1990; Elliott, 1990; Schmaljohn & Patterson, 1990; Elliott et al., 1991).

Bunyamwera virus, which is the prototype of both the family and the Bunyavirus genus, is the only bunyavirus whose three genome RNA segments have been sequenced (Elliott, 1989a, b; Lees et al., 1986). Sequences of the M segments of snowshoe hare (SSH), La Crosse (LAC) (both California serogroup) and Germiston (GER; Bunyamwera serogroup) viruses, and of the S segments of SSH, LAC, Aino (Simbu serogroup), GER and Maguari (MAG; Bunyamwera serogroup) viruses have also been reported (for references see Table 2). Previously we described a six-way comparison of the available bunyavirus N and NSs protein sequences, which highlighted certain conserved regions in the N proteins (Elliott, 1989a). To understand further the relationships between bunyaviruses we have determined the sequences of the S segments of seven other members of this genus. This was facilitated by the development of a one-step
Table 1. Oligonucleotides used in the cloning and mutagenesis of bunyavirus S segment cDNAs

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Position</th>
<th>Sequence*</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUN S+</td>
<td>1 to 15 at 3' end of S segment vRNA</td>
<td>AGTAGGTACTCCAC</td>
<td>Primer for RT-PCR of full-length S cDNA</td>
</tr>
<tr>
<td>BUN S−</td>
<td>Terminal 15 bases of full-length cDNA</td>
<td>AGTAGGTGCTCCAC</td>
<td>Reverse primer for PCR of full-length S cDNA</td>
</tr>
<tr>
<td>GRO S1</td>
<td>39 to 53</td>
<td>gggagaattgatcGAAATTCTTGAGCTTGTT</td>
<td>PCR subcloning of GRO NSs ORF from ACG codon</td>
</tr>
<tr>
<td>GRO S2</td>
<td>55 to 69</td>
<td>gggagaattgatcCTCAAATGCAAACAA</td>
<td>PCR subcloning of GRO NSs ORF from AUG codon</td>
</tr>
<tr>
<td>GRO S3</td>
<td>319 to 304</td>
<td>gggcatgatgactAACCACTTACTTAT</td>
<td>Common downstream PCR primer for GRO cDNA; contains complement of NSs stop codon</td>
</tr>
<tr>
<td>GRO S4</td>
<td>39 to 71</td>
<td>tgtggactGAAATTCTTGAGCTTGCTCAAAGCCCAACAAATGCAACAA</td>
<td>Convert NSs AUG codon to GCC</td>
</tr>
</tbody>
</table>

* Nucleotides corresponding to viral sequences are given in uppercase letters; specific nucleotides mentioned in Purpose column are underlined. Lowercase letters indicate flanking nucleotides that introduce restriction enzyme sites.

Table 2. Summary of bunyavirus S RNA sequences

<table>
<thead>
<tr>
<th>Serogroup</th>
<th>Virus</th>
<th>Total</th>
<th>5' NCR*</th>
<th>3' NCR</th>
<th>A+U (%)</th>
<th>Amino acids</th>
<th>Accession number†</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bunyamwera</td>
<td>BAT</td>
<td>946</td>
<td>68</td>
<td>176</td>
<td>59-5</td>
<td>233 101</td>
<td>X73464</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>CV</td>
<td>950</td>
<td>74</td>
<td>174</td>
<td>59-8</td>
<td>233 101</td>
<td>X73465</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>GER</td>
<td>980</td>
<td>87</td>
<td>191</td>
<td>57-6</td>
<td>233 109</td>
<td>M19420 Gerbaud et al. (1987)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GRO</td>
<td>954</td>
<td>26</td>
<td>226</td>
<td>62-4</td>
<td>233 83</td>
<td>X73466</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>KRI</td>
<td>992</td>
<td>81</td>
<td>209</td>
<td>60-6</td>
<td>233 109</td>
<td>X73467</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>MD</td>
<td>954</td>
<td>76</td>
<td>176</td>
<td>60-0</td>
<td>233 101</td>
<td>X73469</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>NOR</td>
<td>945</td>
<td>74</td>
<td>169</td>
<td>60-5</td>
<td>233 101</td>
<td>X73470</td>
<td>This paper</td>
</tr>
<tr>
<td>California</td>
<td>LAC</td>
<td>981</td>
<td>81</td>
<td>192</td>
<td>58-8</td>
<td>235 92</td>
<td>K00108 Cabradilla et al. (1983)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUM</td>
<td>1077</td>
<td>78</td>
<td>291</td>
<td>58-7</td>
<td>235 97</td>
<td>X73468</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>982</td>
<td>79</td>
<td>195</td>
<td>56-3</td>
<td>235 92</td>
<td>J02390 Bishop et al. (1982)</td>
<td></td>
</tr>
<tr>
<td>Simbu</td>
<td>Aino</td>
<td>850</td>
<td>34</td>
<td>114</td>
<td>55-9</td>
<td>233 91</td>
<td>M22011 Akashi et al. (1984)</td>
<td></td>
</tr>
</tbody>
</table>

* Non-coding region of positive-sense RNA. † Accession no. given in GenBank nucleotide sequence database.

Methods

Viruses. The following bunyaviruses were used in this study: Batai (BAT), Bunyamwera (BUN), Cache Valley (CV), Guaroa (GRO), Kairi (KRI), Lumbo (LUM), Maguari (MAG), Main Drain (MD) and Northway (NOR) viruses. The origins of these viruses and the preparation of virus stocks in baby hamster kidney (BHK) cells were described in Watret et al. (1985).

Viral RNA preparation. BHK cells were infected with the bunyaviruses at an m.o.i. of 1 to 5 p.f.u. per cell and incubated at 31 °C for 2 days. Culture fluids were clarified by centrifugation at 3000 g for 10 min, and virus was pelleted by centrifugation for 20000 g for 2 h in the SW41 rotor. The pellet was resuspended in 400 μl of guanidinium thiocyanate buffer and the RNA extracted with acidic phenol-chloroform (Chomczynski & Sacchi, 1987). The RNA extracted from virions released from 10⁷ infected cells was resuspended in 50 μl of sterile water.

cDNA synthesis and cloning. To synthesize full-length cDNA to the S segment a one-step RT-PCR protocol was used. A 50 μl reaction contained 10 μl RNA, 50 mM-KCl, 20 mM-Tris-HCl pH 8.4, 2.5 mM-MgCl₂, 1 μg/ml BSA, 200 μM dNTPs, 10 pmol of each primer (BUN S+ and BUN S−; Table 1), 20 units RNasin reverse transcriptase (Ambion), 100 units Moloney murine leukaemia virus reverse transcriptase (New England Biolabs) and 1 unit Taq polymerase (Cetus). The reaction mixture was overlaid with mineral oil and incubated at 42 °C for 45 min, and then underwent 30 cycles of 94 °C for 30 s, 54 °C for 30 s and 72 °C for 1 min. The amplified products were purified by electrophoresis through a 1% agarose gel and recovered by binding to a silica matrix (Gene clean, Bio-101). The DNA was blunt-ended and phosphorylated in a single reaction containing...
RT-PCR cloning of bunyavirus S segments

Based on the conservation of the terminal 15 nucleotides in all available S segment sequences of viruses in the Bunyamwera and California serogroups (Elliott et al., 1991) we have developed an RT-PCR protocol to amplify specifically full-length S segment cDNA starting with RNA extracted from unpurified virus (i.e. pelleted virus from infected cell supernatant fluid). Initially we used a two-step procedure, taking an aliquot of a first-strand cDNA reaction primed with oligonucleotide BUN S+ (Table 1) as the template for PCR; subsequently we found that a one-step reaction, incorporating both BUN S+ and BUN S− primers, worked efficiently. Fig. 1 shows the products of such reactions using a variety of bunyavirus RNAs analysed on an agarose gel. The apparently full-length products were isolated from the gel, made blunt-ended and phosphorylated, and ligated to SmaI-digested pUC18 DNA. Nucleotide sequence analysis showed that the majority of the recombinant plasmids contained inserts with complete and correct primer sequences. The primers differ in sequence by one nucleotide (at position 9; Elliott, 1990) but in only one of 30 clones examined was it found that the primer had hybridized at the ‘wrong’ end. Thus, the one-step protocol is an efficient and faithful method of producing full-length bunyavirus S segment cDNA for cloning.

Results and Discussion

Nucleotide sequence determination and analysis

The complete sequence of seven previously undetermined bunyavirus S segments were obtained (Fig. 2) using the described RT-PCR procedure. As a control to monitor the fidelity of our procedure the sequence of MAG S segment was determined and compared to that obtained previously using a conventional reverse transcription approach (Elliott & McGregor, 1989). No differences in the sequences were found between RT-PCR and conventionally produced MAG S cDNA clones. At least three plasmid clones of each S segment cDNA were sequenced; variation between different clones was recorded for only three S segments and in each case the nucleotide present in the majority of the clones was accepted for compiling the sequence shown in Fig. 2. These differences were: in BAT position 55 A or G, G accepted; in CV position 114 A or G, A accepted and position 381 C or T, C accepted; and in LUM position 629 G or A, A accepted.

The sequences of 13 bunyavirus S segments, from viruses in three serogroups of the Bunyavirus genus, are now available and these are summarized in Table 2. The
Fig. 2. For legend see p. 602.
Fig. 2. For legend see p. 602.
segments are generally about 950 to 990 nucleotides long but those of Aino (850 bases) and LUM (1077 bases) are significantly outside this range. By convention the following observations are based on consideration of the positive-sense RNA. The 3' non-coding region (NCR) is always longer than the 5' NCR; GRO virus is notable in having a particularly short 5' NCR and a relatively long 3' NCR compared to the other Bunyamwera serogroup viruses. LUM virus has a similar length 5' NCR to LAC viruses. LUM virus has a similar length 5' NCR to LAC viruses, but its 3' NCR is about 100 bases longer. Bunyavirus S RNAs are richer in A and U residues (55.9 to 62.4%) than G and C residues.

Two ORFs were predicted in each of the seven newly determined sequences and these followed the pattern seen in other bunyavirus S segments. The 5'-most AUG initiated the longer ORF, encoding the N protein, whereas the shorter ORF, encoding NSs, was in the +1 frame though entirely contained within the N protein ORF. A potential third ORF had been reported in the S RNA segments of GER and MAG bunyaviruses, downstream of and in the same frame as the NSs ORF, but the predicted protein product has not been detected in virus-infected cells (Gerbaud et al., 1987; Elliott & McGregor, 1989). A possible third ORF was observed in only the MD and LUM S segments. That in MD was analogous in position to the previously described 75 codon ORF in MAG but was 25 codons shorter; 31 of the 50 residues in the MD ORF were identical to those in MAG, but the significance of this is questionable in view of the constraint imposed by the overlapping N protein ORF in this region. The LUM ORF is in a different reading frame from either the N or NSs ORF and shows essentially no homology with the other predicted ORFs. The status of these ORFs remains uncertain though the available evidence suggests they are not translated during virus infection.

Alignment of the 3' non-coding region revealed the presence of a conserved motif present in all the Bunyamwera and California serogroup virus S RNAs, but not in the S RNA of the Simbu serogroup Aino virus. The sequence alignment (Fig. 3) is presented in the virion (genomic)-sense RNA (i.e. as the 5' end of viral RNA) to show the mRNA transcription termination sites that have been mapped for BUN, GER, LAC and SSH viruses (Patterson & Kolakofsky, 1984; Eshita et al., 1985; Bouloy et al., 1990; Jin & Elliott, 1993). This is a GU-rich area which can be aligned for all viruses and is adjacent to two conserved motifs, U GU CG and CA motifs. Although GRO has both of these motifs the predicted mRNA termination signal is shorter than in the other viruses and does not contain the conserved GGG or GAG triplet upstream of the predicted mRNA termination site. Neither the UGU CA motif nor the CA motif is present in the Aino S RNA nor in the available bunyavirus M and L segment sequences. Thus whether these motifs are involved in mRNA transcription termination or whether their proximity to the termination site is coincidental, and they have other functional significance, is not known.
Comparison of the LUM S sequence against itself using a dot matrix method revealed a line of homology off the main diagonal, suggestive of a repeated sequence in the 3' NCR (Fig. 4a). As seen in Fig. 4(b) (the sequence is written as virion-sense RNA), the region between nucleotides 880 and 937 shows 76% identity with the region from 984 to 1041. Interestingly this region includes the CA motif described above. We do not think that this duplication is an RT-PCR artefact because (i) the identity is not 100%, (ii) independent cDNA clones had the same duplication and (iii) it has not occurred with the other viral RNA templates which have similar sequences. Partial duplication of sequences has not been noted previously as an evolutionary mechanism for bunyaviruses. LUM virus is serologically closely related to Tahyna (TAH) virus and is considered a geographically distinct variety of TAH (Bishop & Shope, 1979); it would be of interest therefore to determine the sequence of the TAH virus S segment to see whether it also contains this duplicated sequence. Evidence that this is not the case comes from Ushijima et al. (1980) who estimated from electrophoretic mobility that the $M_r$ of LUM S RNA is higher than that of TAH virus (0.5 x 10$^6$ compared to 0.45 x 10$^6$).

Identification of GRO virus NSs initiation codon

The 5'-terminal 15 nucleotides of positive-sense bunyavirus S RNAs are conserved but little homology is seen thereafter until the region encompassing the initiation codons for the N and NSs proteins (Fig. 5). (Note that the following description does not apply to the Simbu serogroup virus Aino.) All the viruses except GER and GRO have a tandem AUG at the beginning of the NSs ORF, and for MAG virus either AUG codon can be used to initiate translation of NSs in vitro (A. McGregor & R. M. Elliott, unpublished data). GER NSs protein initiates at a single AUG, positioned equivalent to the second AUG of the other viruses. Between the N and NSs initiation codons there are 16 nucleotides (19 for...
Fig. 5. Nucleotide sequence comparison of the 5' region of positive-sense S segment RNA showing conserved spacing between the initiation codons for the N and NSs proteins. Residues conserved in 11 of the 12 sequences are given in capitals and in the Consensus. Initiation codons are shaded.

GER) and six of these are strictly conserved in all compared sequences. GRO S RNA has UUGACG at the same position as the NSs start site in the other viruses, but has an AUG codon a further 10 nucleotides downstream. None of the other viruses has an AUG codon at this position. The sequence in this part of the GRO RNA was determined directly (see Methods) and was identical to that in the cloned cDNA (data not shown). Hence the question arose whether GRO NSs translation was initiated at the ACG codon which is at the same relative spacing to the N protein AUG as in the other bunyavirus S segments or at the AUG codon further downstream. ACG as a translation initiation codon is rare in eukaryotic systems but has been reported for the Sendai virus C' protein (Curran & Kolakofsky, 1988; Gupta & Patwardhan, 1988) and an adenovirus coat protein (Becerra et al., 1985).

To determine the initiation codon for GRO NSs, messenger-sense RNAs were transcribed in vitro from GRO cDNA cloned into a T7 promoter-containing plasmid and translated in a rabbit reticulocyte lysate; both capped and uncapped transcripts were produced and gave essentially similar results (Fig. 6). Translation of the full-length GRO RNA, derived from plasmid 1, produced N protein (Mr of about 25K) and a major smaller protein with an Mr of about 10K, designated NSs. (The protein encoded by the second ORF in the GRO S segment would have a predicted Mr of about 9.6K.) In radiolabelled infected BHK cell extracts N protein was clearly detected but NSs was not readily identified. However, a 10K protein band which, in contrast to most host proteins, did not diminish in intensity at 48 h post-infection was seen and is a candidate for GRO NSs. Three subgenomic GRO cDNAs which all lacked the AUG initiation codon for N were produced by PCR, and were cloned into plasmids under control of the T7 promoter. Plasmid 2 contained both ACG and AUG codons in the NSs ORF whereas plasmid 3 contained just the AUG codon. In both cases NSs protein was translated from in vitro transcribed RNA, suggesting that the AUG and not the ACG codon was used. This was confirmed with plasmid 4 in which the ACG codon was maintained but the AUG codon was mutated to GCC; in this case no NSs protein was translated from the RNA transcript. Hence the GRO NSs protein (i) is initiated at an AUG which does not follow the spacing from the N protein AUG observed in the other viruses, and (ii) is relatively shorter than the NSs proteins of the other viruses.

**Amino acid sequence comparisons**

The predicted sequences of the N proteins are either 233 amino acids (Bunyamwera and Simbu serogroups) or 235 amino acids (California serogroup) in length whereas the NSs proteins are more heterogeneous, varying in size from 83 to 109 amino acids (Table 2). Comparison of the protein sequences shows the N proteins to be more conserved than the NSs proteins (Table 3). Within a serogroup the N proteins are greater than 62% identical, and at least 40% of amino acids are conserved between serogroups. The NSs proteins show more variation, but again higher identity is seen within a serogroup than between serogroups. (The lower values for GRO NSs are in part because this protein is shorter than the others.) Earlier serological studies had indicated that GRO cross-
reacted with several Bunyamwera serogroup viruses in CF tests, which measure the relatedness of N proteins, but by HI and NT tests (which measure relationships among the M segment-encoded viral glycoproteins) such relationships were not evident, and cross-reactivity with California serogroup viruses was noted (Bishop & Shope, 1979). Our sequence data clearly confirm that GRO has a Bunyamwera serogroup-like rather than California serogroup-like N protein (Tables 2 and 3).

Alignments of the 13 available N and NSs protein sequences are shown in Fig. 7. The overall homology observed between N proteins is localized to certain regions e.g. residues 39 to 43, 90 to 116 and 125 to 165 (Fig. 7a). These regions may elicit the CF antibodies that cross-react throughout the Bunyavirus genus. Inspection of the alignment indicates that between these regions of global homology there are areas conserved within a serogroup e.g. residues 115 to 124. Delineation of these conserved regions suggests targets for site-directed mutagenesis to investigate specific functions of the N protein such as interaction with the viral RNA or with the L protein. Little global homology is seen among the NSs proteins (Fig. 7b). Near the carboxy terminus the motif TGTxQ is conserved in Bunyamwera and California serogroup viruses, and the tripeptide LPS is conserved in all viruses except GRO. Since the function of the bunyavirus NSs protein remains unknown the significance of these motifs is obscure. Variation in the length of the NSs proteins occurs at the carboxy terminus, with the exception of GRO NSs, which is also truncated at the amino terminus. Bunyamwera serogroup viruses except GRO have a conserved region in their carboxy-terminal extension, and other serogroup-specific conserved regions can be observed throughout the proteins. Notably, however, GRO NSs is more divergent in these areas than the NSs proteins of the other Bunyamwera serogroup viruses.

### Table 3. Amino acid identities between bunyavirus N proteins and between bunyavirus NSs proteins

<table>
<thead>
<tr>
<th>Virus</th>
<th>BAT</th>
<th>BUN</th>
<th>CV</th>
<th>GER</th>
<th>GRO</th>
<th>KRI</th>
<th>MAG</th>
<th>MD</th>
<th>NOR</th>
<th>LAC</th>
<th>LUM</th>
<th>SSH</th>
<th>Aino</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAT</td>
<td>91.8</td>
<td>94.0</td>
<td>77.3</td>
<td>70.4</td>
<td>68.2</td>
<td>93.1</td>
<td>89.3</td>
<td>95.1</td>
<td>43.4</td>
<td>42.5</td>
<td>45.1</td>
<td>41.7</td>
<td></td>
</tr>
<tr>
<td>BUN</td>
<td>89.0</td>
<td>90.6</td>
<td>75.1</td>
<td>68.7</td>
<td>70.8</td>
<td>91.0</td>
<td>90.1</td>
<td>92.7</td>
<td>43.4</td>
<td>43.4</td>
<td>45.1</td>
<td>43.0</td>
<td></td>
</tr>
<tr>
<td>CV</td>
<td>94.1</td>
<td>88.1</td>
<td>75.1</td>
<td>70.8</td>
<td>69.1</td>
<td>95.7</td>
<td>89.7</td>
<td>96.1</td>
<td>44.6</td>
<td>44.2</td>
<td>45.9</td>
<td>42.2</td>
<td></td>
</tr>
<tr>
<td>GER</td>
<td>68.7</td>
<td>71.0</td>
<td>71.7</td>
<td>63.5</td>
<td>62.2</td>
<td>73.8</td>
<td>74.3</td>
<td>76.0</td>
<td>45.5</td>
<td>42.1</td>
<td>46.4</td>
<td>43.2</td>
<td></td>
</tr>
<tr>
<td>GRO</td>
<td>43.9</td>
<td>45.8</td>
<td>43.9</td>
<td>43.2</td>
<td>69.5</td>
<td>69.1</td>
<td>68.7</td>
<td>70.0</td>
<td>47.6</td>
<td>45.1</td>
<td>47.6</td>
<td>44.2</td>
<td></td>
</tr>
<tr>
<td>KRI</td>
<td>67.3</td>
<td>67.3</td>
<td>69.3</td>
<td>56.5</td>
<td>51.2</td>
<td>67.8</td>
<td>70.4</td>
<td>69.5</td>
<td>48.1</td>
<td>45.5</td>
<td>49.7</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>MAG</td>
<td>85.1</td>
<td>89.0</td>
<td>97.0</td>
<td>69.7</td>
<td>42.7</td>
<td>68.3</td>
<td>91.4</td>
<td>94.9</td>
<td>43.4</td>
<td>42.9</td>
<td>44.6</td>
<td>41.7</td>
<td></td>
</tr>
<tr>
<td>MD</td>
<td>90.1</td>
<td>86.1</td>
<td>94.1</td>
<td>67.7</td>
<td>47.6</td>
<td>69.3</td>
<td>93.1</td>
<td>91.9</td>
<td>43.4</td>
<td>41.6</td>
<td>44.6</td>
<td>40.4</td>
<td></td>
</tr>
<tr>
<td>NOR</td>
<td>94.1</td>
<td>88.1</td>
<td>96.0</td>
<td>68.7</td>
<td>45.1</td>
<td>70.3</td>
<td>95.1</td>
<td>95.1</td>
<td>44.2</td>
<td>42.5</td>
<td>45.9</td>
<td>42.1</td>
<td></td>
</tr>
<tr>
<td>LAC</td>
<td>38.2</td>
<td>38.2</td>
<td>38.9</td>
<td>34.1</td>
<td>35.5</td>
<td>32.7</td>
<td>39.3</td>
<td>40.5</td>
<td>39.3</td>
<td>84.7</td>
<td>90.6</td>
<td>45.0</td>
<td></td>
</tr>
<tr>
<td>LUM</td>
<td>34.0</td>
<td>30.9</td>
<td>41.1</td>
<td>32.3</td>
<td>33.3</td>
<td>34.0</td>
<td>41.1</td>
<td>44.4</td>
<td>41.1</td>
<td>75.0</td>
<td>87.7</td>
<td>44.6</td>
<td></td>
</tr>
<tr>
<td>SSH</td>
<td>31.9</td>
<td>29.7</td>
<td>31.5</td>
<td>31.9</td>
<td>36.4</td>
<td>33.7</td>
<td>31.9</td>
<td>37.1</td>
<td>31.9</td>
<td>85.8</td>
<td>76.1</td>
<td>45.0</td>
<td></td>
</tr>
<tr>
<td>Aino</td>
<td>28.9</td>
<td>30.7</td>
<td>28.6</td>
<td>24.2</td>
<td>38.0</td>
<td>31.9</td>
<td>27.8</td>
<td>31.1</td>
<td>30.0</td>
<td>30.7</td>
<td>37.1</td>
<td>34.1</td>
<td></td>
</tr>
</tbody>
</table>

Amino acid identity between NSs proteins (%)
Fig. 7. Alignment of (a) N proteins and (b) NSs proteins of 13 bunyaviruses. The alignments were generated using the PILEUP and PRETTY programs in the UWCGG package. The consensus sequence is based on conservation of amino acid residue in at least 12 of 13 sequences.
Conserlsus

Fig. 7. Cont.

81
tmvLPStd...svdtpgtylrky.................
MAG
tmvLPStd...svdtpgtylrkf................
NOR
tmvLPStd...svdtpgtylrry.............
BAT
tmvLPStd...svdtpgtylrrf.............
MD
tmvLPStd...svdtpgtylrky.............
BUN
tmvLPStd...svdtpgtylrrc.............
KRI
mmvLPStv...svdllpgtcteyllqenqn.
GER
ttvLPStv...svdtpgtylestlqrqmnkss
Gro
mhshfie...........................
LAC
ttiLPStd...ygli.......................
SSH
atitLPStd...cqdi.......................
LUM
ttiLPStg...cqgiwldgc.................
AINO
cqtLPslsivsqdl---------------------

Evolutionary relationships of bunyavirus S segments

Viruses within the Bunyavirus genus have been categorized into serogroups generally on the pattern of reactivity in the CF test. Comparison of the 13 available bunyavirus N protein sequences shows that their classification into three serogroups correlates with the relatedness of their N proteins, as illustrated diagrammatically in Fig. 8. Within the Bunyamwera serogroup several complexes are recognized, and by CF tests BAT, BUN, CV, MAG, MD and NOR were virtually indistinguishable whereas GER, GRO and KRI were distinct (summarized by Bishop & Shope, 1979). Again the sequence data are in complete agreement (Fig. 8). However, in NT tests MD was in a distinct complex and GER fell into the Bunyamwera complex (Hunt & Calisher, 1979); these data thus suggest that genome segment reassortment has occurred in nature during the evolution of these viruses in the Bunyamwera complex. Furthermore GRO also appears to be the result of an ancestral reassortment event, but in this case between parental viruses which would now be classified into different serogroups. Interestingly, comprehensive attempts to demonstrate experimentally any reassortment between present members of the Bunyamwera and California serogroups have been unsuccessful (Elliott et al., 1984). The molecular basis for the restriction of genome segment reassortment remains to be elucidated.

Phylogenetic analyses of bunyavirus S segment sequences, using programs in PHYLIP (Felsenstein, 1991), gave results which largely mimic the dendrograms shown in Fig. 8, especially with regard to the major nodes. However it appeared that BAT was more closely related to NOR, and CV was more closely related to MAG (data not shown). Sequences of more bunyavirus isolates, particularly representatives of the other serogroups, will be required to gain a fuller picture of bunyavirus evolution. The methodology we have described above would facilitate such studies.

We thank Anne Bridgen for thoughtful comments on the manuscript. D.C.P. was in receipt of a Glaxo Group Research Scholarship. R.M.E. is an MRC Senior Fellow and this work was supported by an MRC Project Grant.

References


(Received 27 September 1993; Accepted 22 October 1993)