Complete sequence of Great Island virus and comparison with the T2 and outer-capsid proteins of Kemerovo, Lipovnik and Tribec viruses (genus *Orbivirus*, family *Reoviridae*)

Mourad Belhouchet,1,† Fauziah Mohd Jaafar,1,† Robert Tesh,2 Jonathan Grimes,3 Sushila Maan,1 Peter P. C. Mertens1 and Houssam Attoui1

1Department of Vector-Borne Diseases, Institute for Animal Health, Ash Road, Pirbright GU24 0NF, UK
2Department of Pathology, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555–0609, USA
3Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, Oxford OX3 7BN, UK

The complete nucleotide sequence of Great Island virus (GIV) genome was determined, along with genome segments (Seg) 1, 2 and 6 of Kemerovo (KEMV), Lipovnik (LIPV) and Tribec (TRBV) viruses. All four viruses, together with Broadhaven virus, are currently classified within the species *Great Island virus* and have been isolated from ticks, birds or humans. Sequence comparisons showed that Seg-4 of GIV encoded the outer-capsid protein responsible for cell attachment, although it was approximately half the length of its counterpart in the *Culicoides* or mosquito-transmitted orbiviruses. A second overlapping ORF (in the +2 reading frame) was identified in Seg-9 of GIV, encoding a putative dsRNA-binding protein. Phylogenetic analyses of the RNA-dependent RNA polymerase (Pol) and T2 protein amino acid sequences indicated that the tick-borne orbiviruses represent an ancestral group from which the mosquito-borne orbiviruses have evolved. This mirrors the evolutionary relationships between the arthropod vectors of these viruses, supporting a co-speciation hypothesis for these arboviruses and their arthropod-vectors. Phylogenetic analyses of the T2 proteins of KEMV, LIPV, TRBV and GIV (showing 82% amino acid identity) correlated with the early classification of Great Island viruses as two distinct serocomplexes (Great Island and Kemerovo serocomplexes). Amino acid identity levels in the VP1(Pol) and T2 proteins between the two serocomplexes were 73 and 82%, respectively, whilst those between previously characterized *Orbivirus* species are 53–73% and 26–83%, respectively. These data suggest that, despite limited genome segment reassortment between these two groups, their current classification within the same *Orbivirus* species could be re-evaluated.

**INTRODUCTION**

The genus *Orbivirus* contains 22 virus species that are vectored by *Culicoides* midges, ticks, phlebotomine flies and anopheline or culicine mosquitoes. Several tick-borne orbiviruses can infect humans, and these include Kemerovo virus (KEMV), Lipovnik virus (LIPV) and Tribec virus (TRBV), which are members of the species *Great Island virus*. Currently, there are 36 named serotypes/strains within this species, most of which have been isolated from seabirds or ticks, including *Argas*, *Ornithodoros* and *Ixodes* species. The current classification of these isolates within a single species relies primarily on reassortment and limited serological studies (Mertens et al., 2005).

The prototype ‘Great Island virus’ was first isolated from engorged nymphs of *Ixodes uriae* collected on 27 July 1971 from puffin (*Fratercula arctica*) burrows on Great Island, Newfoundland, Canada. Antibodies against Great Island virus (GIV) have been detected in sera from Atlantic...
puffins (F. arctica) and Leach’s petrel (Oceanodroma leucorhoa) (Karabatsos, 1985).

KEMV was first isolated from female imagoes of *Ixodes persulcatus* collected on 21 May 1962 in the Kemerovo region of Russia. KEMV (L75 strain) was subsequently isolated from human patients with aseptic meningitis (Libikova et al., 1970), and seroconversion rates of 2.8% for KEMV were detected in healthy humans in the Kemerovo region of western Siberia.

LIPV was first isolated from female imagoes of *Ixodes ricinus* collected between 25 and 30 May 1963 in Lipovnik village, Slovakia, whilst TRBV was isolated from nymphs of *I. ricinus* collected on 25 April 1963 in the Tribeč Mountains in Slovakia. TRBV was also isolated from the red-back mouse (*Clethrionomys glareolus*), pine mouse (*Pitymys subterraneus*) and goats. Seroconversion rates for TRBV and LIPV of 3 and 18%, respectively, were detected in apparently healthy humans in western and eastern Slovakia. These sites represent natural foci of infection for these viruses (Gresíkova et al., 1966; Libikova et al., 1970).

Although sequence data have been generated for several of the insect-borne orbiviruses [such as the *Culicoides*-borne bluetongue virus (BTV), African horse sickness virus (AHSV) and epizootic hemorrhagic disease virus (EHDV) (Mertens et al., 2005), and mosquito-borne Yunnan orbivirus (YUOV) and Peruvian horse sickness virus (PHSV)], only three of the tick-borne orbiviruses have been sequenced to date. These are St Croix River virus (SCRV, full sequence; Attoui et al., 2001), Broadhaven virus (BRDV, partial sequence; Moss et al., 1992) and Sandy Bay virus (SBAV, partial sequence; Major et al., 2009). Previous comparisons of homologous proteins of the insect- and tick-borne orbiviruses have shown only 23–38% amino acid identity, revealing high levels of genetic diversity within the genus Orbivirus.

This paper reports the full-length sequence analysis of the GIV genome, and comparisons of the genome segments (Seg) 1, 2 and 6 encoding the RNA-dependent RNA polymerase (Pol) and the subcore shell protein (T2), respectively, of KEMV, LIPV and TRBV.

RESULTS

**Virus propagation, electrophoretotype analysis and electron microscopy**

GIV, KEMV, LIPV and TRBV particles were purified from infected BHK-21 cells using caesium chloride gradient centrifugation. Electron microscopy showed particles with a mean diameter of 55 nm, showing defined ring-shaped capsomeres characteristic of orbivirus core particles (Attoui et al., 2001, 2005) (see Supplementary Fig. S1a, available in JGV Online).

Genomic RNAs extracted from purified viruses gave an almost identical migration pattern on agarose gels, except that Seg-9 and -10 of GIV migrated slightly faster than those of the other three viruses (see Supplementary Fig. S1b).

**General sequence analysis**

The ten dsRNA segments of the GIV genome were converted into full-length cDNAs, cloned and sequenced. Genome segments 1, 2 and 6 of KEMV, LIPV and TRBV were also cloned and sequenced. The lengths of the genome segments and the proteins they encode are given in Table 1. Analysis of the 5’- and 3’-terminal non-coding regions showed that all of the GIV segments shared five fully conserved nucleotides at their 5’ ends and two fully conserved nucleotides at their 3’ ends (5’-GUAAAG/UA/UCG-3’/CGGUCC-3’). Moreover, the first two and last two nucleotides of all segments were inverted complements and were identical to those found in other orbiviruses.

Most of the GIV genome segments contained a single major ORF, which spanned almost the entire length of the positive strand. The only exception was Seg-9, which contained two overlapping out-of-phase ORFs. The first (upstream) ‘ORF-1’ (+1 reading frame, nt 55–1017) encoded the viral helicase VP6(Hel) and was a homologue of VP6 of orbiviruses. The second and smaller ‘ORF-2’ (+2 reading frame, nt 174–745) encoded a novel 190 aa (22.5 kDa) protein. The initiation codon of ORF-1 (TCCATGT), like that of BTV Seg-9, did not have a strong Kozak context (Kozak, 1987). The Kozak context of the ORF-2 initiation codon (AGAATGA) was also only ‘moderate’. However, many ORFs carried by genome segments of different reoviruses also have ‘weak or moderate’ contexts but still appear to be expressed and translated efficiently in infected cells. Analysis of the amino acid sequence of the novel 190 aa protein using the Pfam programme (http://pfam.sanger.ac.uk/search) showed a 72 aa fragment (aa 82–153) with 39% similarity to dsRNA-binding domains of similar length (~68 aa) in other reoviruses. On the basis of its putative role as a dsRNA-binding protein, we identified this protein as ‘VP6(dBP)’.

Comparisons of GIV amino acid sequences with those of BRDV, SBAV, KEMV, TRBV and LIPV (see Supplementary Tables S1 and S2, available in JGV Online) showed identity values that ranged from 67 to 91.8%. The only GIV protein that did not show a significant level of identity with other orbivirus proteins was VP6(dBP) (Table 2).

**Comparison of GIV VP1(Pol) with the polymerases of other orbiviruses**

Identity levels of 53–73% were detected in a previous study (Attoui et al., 2002) between different insect-transmitted Orbivirus species, including AHSV, EHDV, BTV, equine encephalosis virus and Palyam virus (PALV). In contrast only ~35% amino acid identity was detected between these insect-transmitted viruses and the tick-borne SCRV (Attoui et al., 2001).
Fig. 1 shows a neighbour-joining tree of the polymerase amino acid sequences from GIV, KEMV, LIPV and TRBV aligned with those of other Orbivirus species and representative members of other genera within the family Reoviridae. The tree indicates that these tick-borne viruses provide a ‘root’ for all of the insect-borne orbiviruses. VP1(Pol) of GIV showed only 45–47 % amino acid identity to the polymerase of the insect-borne orbiviruses, and only 41 % to SCRV. In contrast, identity levels of 73 % were detected between VP1(Pol) of GIV and that of KEMV, TRBV or LIPV, whilst 80 % identity was detected between KEMV and TRBV or LIPV, and 99 % identity was detected between LIPV and TRBV. The higher values among KEMV, LIPV and TRBV, and the lower values between GIV and KEMV, LIPV or TRBV, correlated with the early classification of KEMV, LIPV and TRBV in a separate serocomplex from GIV.

Comparisons of T2 subcore proteins

The minimum levels of amino acid identity for the T2 protein within individual virus species are 99 % for AHSV, 99 % for Warrego virus (WARV), 98 % for PALV, 95.5 % for Wongorr virus, 95 % for EHDV; 94 % for Corriparta virus; 92.3 % for Wallal virus (WALV) and 86 % for BTV (Attoui et al., 2001). However, the T2 protein shows lower levels of identity (26–83 %) between distinct Orbivirus species (the highest level being between isolates of BTV and EHDV).

The T2 protein of the Culicoides-borne orbiviruses, including BTV, AHSV, EHDV, WALV, Eubenangee virus, WARV and PALV, is encoded by Seg-3 \[\text{VP3(T2)}\]. In contrast, the equivalent proteins [\text{VP2(T2)}] of the more distantly related SCRV and BRDV are encoded by Seg-2 and are part of a separate phylogenetic cluster.

### Table 1. Lengths of dsRNA segments, encoded putative proteins and 5’ and 3’ non-coding regions

<table>
<thead>
<tr>
<th>Virus</th>
<th>Accession no.</th>
<th>Segment length (bp)</th>
<th>Protein Length (aa)</th>
<th>Mass (Da)*</th>
<th>5’NCR Terminal</th>
<th>3’NCR Terminal</th>
<th>Mol% G+C</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIV</td>
<td>Seg-1</td>
<td>HM543465 3897</td>
<td>1285</td>
<td>146 850</td>
<td>11 5’-GUAAUA</td>
<td>AUCCUAC-3’</td>
<td>31 56</td>
</tr>
<tr>
<td></td>
<td>Seg-2</td>
<td>HM543466 2794</td>
<td>908</td>
<td>102 908</td>
<td>18 5’-GUAAAC</td>
<td>AAGAUAC-3’</td>
<td>52 58</td>
</tr>
<tr>
<td></td>
<td>Seg-3</td>
<td>HM543467 1936</td>
<td>635</td>
<td>72 841</td>
<td>5 5’-GUAAUA</td>
<td>AAGCUAC-3’</td>
<td>26 57</td>
</tr>
<tr>
<td></td>
<td>Seg-4</td>
<td>HM543468 1731</td>
<td>531</td>
<td>59 867</td>
<td>39 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>98 59</td>
</tr>
<tr>
<td></td>
<td>Seg-5</td>
<td>HM543469 1722</td>
<td>551</td>
<td>62 326</td>
<td>17 5’-GUAAA</td>
<td>AGGAUAC-3’</td>
<td>52 59</td>
</tr>
<tr>
<td></td>
<td>Seg-6</td>
<td>HM543470 1666</td>
<td>537</td>
<td>59 512</td>
<td>20 5’-GUAAA</td>
<td>GUCCUAC-3’</td>
<td>35 59</td>
</tr>
<tr>
<td></td>
<td>Seg-7</td>
<td>HM543471 1181</td>
<td>357</td>
<td>39 643</td>
<td>17 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>93 59</td>
</tr>
<tr>
<td></td>
<td>Seg-8</td>
<td>HM543472 1172</td>
<td>359</td>
<td>38 876</td>
<td>45 5’-GUAAA</td>
<td>AGGAUAC-3’</td>
<td>47 59</td>
</tr>
<tr>
<td></td>
<td>Seg-9</td>
<td>HM543473 1056</td>
<td>321/190</td>
<td>34 452/22 527</td>
<td>54 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>39 58</td>
</tr>
<tr>
<td></td>
<td>Seg-10</td>
<td>HM543474 703</td>
<td>171</td>
<td>19 404</td>
<td>145 5’-GUAAA</td>
<td>AGGAUAC-3’</td>
<td>45 58</td>
</tr>
<tr>
<td>Consensus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5’-GUAAA$^{5’G/C}$$^{3’G/A/C}$</td>
<td>$^{5’G/C}$</td>
<td>$^{3’G/A/C}$</td>
</tr>
<tr>
<td>BRDV</td>
<td>Seg-2</td>
<td>M87875 2792</td>
<td>908</td>
<td>102 988</td>
<td>18 5’-GUAAAC</td>
<td>AAGAUAC-3’</td>
<td>50 56</td>
</tr>
<tr>
<td></td>
<td>Seg-4</td>
<td>X82599 1714</td>
<td>537</td>
<td>60 545</td>
<td>38 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>65 59</td>
</tr>
<tr>
<td></td>
<td>Seg-5</td>
<td>M58030 1658</td>
<td>480</td>
<td>52 524</td>
<td>21 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>198 59</td>
</tr>
<tr>
<td></td>
<td>Seg-7</td>
<td>M87876 1174</td>
<td>356</td>
<td>40 199</td>
<td>18 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>89 56</td>
</tr>
<tr>
<td></td>
<td>Seg-10</td>
<td>M83197 702</td>
<td>205</td>
<td>22 044</td>
<td>17 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>70 56</td>
</tr>
<tr>
<td>LIPV</td>
<td>Seg-1</td>
<td>HM543475 3892</td>
<td>1284</td>
<td>145 799</td>
<td>10 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>30 56</td>
</tr>
<tr>
<td></td>
<td>Seg-2</td>
<td>HM543476 2793</td>
<td>908</td>
<td>102 705</td>
<td>18 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>51 56</td>
</tr>
<tr>
<td></td>
<td>Seg-6</td>
<td>HM543477 1509†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>TRBV</td>
<td>Seg-1</td>
<td>HM543478 3892</td>
<td>1284</td>
<td>145 598</td>
<td>10 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>30 56</td>
</tr>
<tr>
<td></td>
<td>Seg-2</td>
<td>HM543479 2793</td>
<td>908</td>
<td>102 757</td>
<td>18 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>51 58</td>
</tr>
<tr>
<td></td>
<td>Seg-6</td>
<td>HM543480 1668</td>
<td>537</td>
<td>59 555</td>
<td>23 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>34 57</td>
</tr>
<tr>
<td>KEMV</td>
<td>Seg-1</td>
<td>HM543481 3896</td>
<td>1285</td>
<td>146 028</td>
<td>10 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>31 55</td>
</tr>
<tr>
<td></td>
<td>Seg-2</td>
<td>HM543482 2972</td>
<td>908</td>
<td>102 750</td>
<td>18 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>50 57</td>
</tr>
<tr>
<td></td>
<td>Seg-6</td>
<td>HM543483 1667</td>
<td>537</td>
<td>59 444</td>
<td>22 5’-GUAAA</td>
<td>AAGAUAC-3’</td>
<td>34 57</td>
</tr>
</tbody>
</table>

*Theoretical molecular mass calculated from deduced amino acid sequences.
†Partial sequence.
Table 2. Correspondence between GIV, BTV (a typical insect-borne orbivirus) and SCRV (a tick-borne orbivirus belonging to a distinct species)

For functions and abbreviations (shown in parentheses) used to indicate these roles, see Mertens et al. (2005). NM, No match.

<table>
<thead>
<tr>
<th>GIV</th>
<th>SCRV (% aa identity)</th>
<th>BTV (% aa identity)</th>
<th>YUOV (% aa identity)</th>
<th>Putative function*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seg-1, VP1(Pol)</td>
<td>Seg-1, VP1(Pol) (41)</td>
<td>Seg-1, VP1(Pol) (47)</td>
<td>Seg-1, VP1(Pol) (47)</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>Seg-2, VP2(T2)</td>
<td>Seg-2, VP2(T2) (25)</td>
<td>Seg-3, VP3(T2) (36)</td>
<td>Seg-2, VP2(T2) (46)</td>
<td>T2, major subcore Protein</td>
</tr>
<tr>
<td>Seg-3, VP3(Cap)</td>
<td>Seg-4, VP4(Cap) (37)</td>
<td>Seg-4, VP4(Cap) (45)</td>
<td>Seg-4, VP4(Cap) (38)</td>
<td>Minor core and capping enzyme (Cap)</td>
</tr>
<tr>
<td>Seg-4, NS1</td>
<td>Seg-6, NS1 (22)</td>
<td>Seg-5, NS1 (24)</td>
<td>Seg-5, NS1 (26)</td>
<td>Tubules (TuP)</td>
</tr>
<tr>
<td>Seg-5, VP4</td>
<td>Seg-3, VP3 (20)</td>
<td>Seg-2, VP2 (21)</td>
<td>Seg-3, VP3 (21)</td>
<td>Similar to outer-shell protein VP2 of BTV, neutralization epitope</td>
</tr>
<tr>
<td>Seg-6, VP5</td>
<td>Seg-5, VP5 (36)</td>
<td>Seg-6, VP5 (32)</td>
<td>Seg-6, VP5 (39)</td>
<td>VP5, outer-capsid Protein</td>
</tr>
<tr>
<td>Seg-7, VP7</td>
<td>Seg-8, VP7 (25)</td>
<td>Seg-7, VP7 (23)</td>
<td>Seg-8, VP7 (31)</td>
<td>Major core surface protein, T13 (780 copies)</td>
</tr>
<tr>
<td>Seg-8, NS2</td>
<td>Seg-8, NS2 (18)</td>
<td>Seg-8, NS2 (23)</td>
<td>Seg-7, NS2 (33)</td>
<td>Viral inclusion bodies (VIP)</td>
</tr>
<tr>
<td>Seg-9, VP6</td>
<td>Seg-9, VP6 (42)</td>
<td>Seg-9, VP6 (23)</td>
<td>Seg-9, VP6 (28)</td>
<td>Minor core protein, helicase (Hel)</td>
</tr>
<tr>
<td>Seg-10, NS3</td>
<td>Seg-10, NS3 (23)</td>
<td>Seg-10, NS3 (24)</td>
<td>Seg-10, NS3 (20)</td>
<td>dsRNA-binding</td>
</tr>
<tr>
<td>Seg-9, VP6(dBP)</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>Non-structural (virus release)</td>
</tr>
</tbody>
</table>

*Putative functions of GIV proteins by comparison with the previously established functions of BTV.

(Vig. 2) that includes the mosquito-borne species Corriparta virus and Wongorr virus. VP2 (the second largest protein) of GIV, KEMV, LIPV and TRBV was identified as the T2 subcore capsid shell protein by comparisons with VP2(T2) of BRDV and SCRV (Attoui et al., 2001; Moss & Nuttall, 1994) and VP3(T2) of BTV (Grimes et al., 1998). Phylogenetic analysis indicated that VP2 of GIV, KEMV, LIPV and TRBV clustered together with other members of the VP2(T2) group.

VP2(T2) of GIV showed 45.6–56.2% amino acid identity to that of the mosquito-borne orbiviruses, 35.2–41.6% identity to the Culicoides-borne orbiviruses and 23.7% identity to SCRV. An amino acid identity level of 91.8% was detected with BRDV, indicating that GIV and BRDV belong to the same species. The level of amino acid identity between LIPV and TRBV or KEMV was also high (at 91.1 and 99.4%, respectively), indicating that these three viruses also belong to a single species. However, a lower level of amino acid identity (82%) was detected in the T2 protein between GIV and LIPV, TRBV or KEMV. This value correlates with the early classification of these viruses within two distinct serocomplexes, the Great Island and Kemerovo complexes. These data suggest that the species Great Island virus is on the ‘borderline’ of classification as two distinct species.

The shape parameter $\alpha$, measuring variability rates at different sites, was calculated as 1.959 for the available orbivirus T2 protein sequences, including those of GIV, LIPV, TRBV and KEMV. An $\alpha$ value $>1$ indicates that most sites have similar rates of variability. When $\alpha \leq 1$ (there is a relatively high level of rate variation), many sites have very low rates, but there are evolutionary hot spots with higher rates. The calculated value was in agreement with different sites evolving at similar rates within the T2 protein.

Comparison of outer-capsid proteins

The cell-attachment proteins of the orbiviruses are highly variable, even among different serotypes within a single species (Attoui et al., 2009; Maan et al., 2007). Phylogenetic comparisons indicated that outer-capsid protein VP4 of GIV (encoded by Seg-5) was functionally equivalent to the outermost capsid protein VP2 of Culicoides-borne orbiviruses and VP3 of mosquito-borne orbiviruses or SCRV (Table 2). VP4 of GIV is therefore considered likely to be the cell-attachment protein, which represents a target for neutralizing antibodies.

The other outer-capsid protein of GIV is VP5 and is equivalent to VP5 of the insect-borne orbiviruses and SCRV (Table 2). VP5 can also be variable among different serotypes within a single Orbivirus species, but to a lesser extent than VP2 (Attoui et al., 2009; Maan et al. 2007; Singh et al., 2004). Sequence comparisons of VP5 from GIV (determined in this study), BRDV and SbAV (Major et al., 2009; Moss et al., 1990) have revealed three single-base deletions within VP5 of BRDV relative to the GIV.
sequence (at positions 151, 157 and 215) and three zones of multiple deletions (at positions 419–445, 472–492 and 525–530). Amino acid identity levels between VP5 of GIV, KEMV, LIPV, TRBV, BRDV and SBaV ranged from 60.2 to 95.8 % (see Supplementary Table S1). These analyses showed that, although VP5 of GIV represents a ‘root’ for the homologous protein from KEMV, LIPV, TRBV and SBaV, BRDV is located at the root of all of the isolates compared (Fig. 3).

The outermost and ‘cell-attachment’ protein (identified as VP2 or VP3) of the insect-borne orbiviruses is almost twice as long as its counterpart in members of the tick-borne Great Island virus species. A local BLAST search indicated that VP4 of GIV is related to the C-terminal half (aa 483–954) of VP2 from BTV, EHDV and ASHV (Culicoides transmitted) and VP3 from YUOV and PHSV (both mosquito transmitted), with 28–30 % amino acid sequence identity.

DISCUSSION

The full-genome sequence of GIV is reported here, along with partial sequences for the zoonotic viruses KEMV, LIPV and TRBV. These data, which will facilitate the development of sequence-specific RT-PCR assays for epidemiological studies of the GIV group, were used to analyse relationships with other previously characterized orbiviruses.

Unlike the other nine genome segments, which are monocistronic, Seg-9 of GIV has two overlapping but out-of-phase ORFs. The first, ‘ORF-1’, encodes the helicase, VP6(Hel), thought to be involved in RNA replication. The second, ‘ORF-2’, encodes a novel protein designated VP6(dBP), which contains a dsRNA-binding domain similar to those located in dsRNA-binding proteins of other dsRNA viruses (e.g. VP12 of Banna virus; Attoui et al., 2000a; Mohd Jaafar et al., 2005). The presence of these
dsRNA-binding motifs suggests that, like α3 of the mammalian orthoreoviruses, VP6(dBP) might be involved in counteracting innate immunity within infected cells, particularly activation of protein kinase R and RNA interference (Beattie et al., 1995; Csorba et al., 2007). Moreover, because of the suggested activity of VP6(dBP),

Fig. 2. Phylogenetic comparison of the T2 proteins (the major component of the subcore shell) of GIV, KEMV, TRBV, LIPV and other orbivirus species. This protein is equivalent to the VP3(T2) protein of BTV, the prototype Orbivirus species, and to the VP2(T2) of two tick-borne orbiviruses: SCRV and BRDV. As many of the available sequences are incomplete, the analysis (presented as a radial tree) is based on partial sequences (aa 393–548 relative to the BTV-10 sequence). GenBank accession numbers, abbreviations and further details of the sequences and viruses used are included in Supplementary Table S4 (available in JGV Online). The tree was constructed using the P-distance algorithm. Values at the nodes indicate bootstrap confidence. Bar, number of substitutions per site.

Fig. 3. Phylogenetic tree built with the VP5 proteins of six isolates from the currently recognized Great Island virus species. The tree shows that GIV roots the other isolates except for BRDV. BRDV has been isolated as a temperature-sensitive mutant of a tick-borne orbivirus (FT363; Moss & Nuttall, 1986). Being a temperature-sensitive mutant could explain the important deletions identified in Seg-5 of BRDV, hence its position rooting all other isolates. Bar, number of substitutions per site. GenBank accession numbers are indicated.
this protein might also be involved in counteracting the innate antiviral response to dsRNA provided by retinoic acid inducible gene I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5) (Andrejeva et al., 2004; Kang et al., 2002; Yoneyama, et al., 2004, 2005).

The orbivirus outer-capsid and cell-attachment protein, which is the primary determinant of virus serotype, is represented by VP3 of SCRV (tick-borne), VP2 of Culicoides-borne orbiviruses and VP3 of mosquito-borne orbiviruses such as YUOV and PHSV. However, the functionally equivalent protein of orbiviruses classified within the Great Island virus species, which are tick-borne, is significantly smaller (approximately half the size) and is identified as VP4. Sequence alignments showed that VP4 of GIV matches the C-terminal half of outer-capsid protein VP2 or VP3 of the insect-borne orbiviruses. It is therefore not surprising that VP4 of GIV has only a low level of identity to the homologous outer-capsid protein of the other orbiviruses (e.g. ~10 % with VP2 of BTV-10).

The insect- and tick-borne orbivirus protein forming the T13 outer layer of the core (VP7) is more highly conserved than the outer-capsid proteins. It is also immunodominant and represents the primary antigenic determinant of virus serogroup (or species) (Mertens et al., 2005). Sequence comparisons showed that VP7(T13) of GIV exhibited a high level of amino acid identity with that of BRDV (80 %), but a much lower level of identity with its homologues from SCRV (25 %), BTV (24 %) and YUOV (31 %). This is consistent with the classification of GIV within the same orbivirus serogroup/species as BRDV.

Previous studies have shown 36–73 % amino acid identity between the polymerase proteins of different Orbivirus species (Attoui et al., 2001). The lowest value was between SCRV and BTV, whilst the highest values were between BTV and EHDV, which are regarded as very closely related species with some serological cross-reactions (Attoui et al., 2001). The polymerase of GIV (VP1) showed between 41 and 47 % amino acid identity to the RNA polymerase of the other characterized orbiviruses, confirming its classification as a distinct species in the genus Orbivirus.

Levels of identity in the subcore shell ‘T2’ protein between Orbivirus species showed a greater range (26–83 %) than between the polymerase proteins. However, as with the polymerase, the lowest value previously detected was between SCRV and BTV, whilst the highest was again between BTV and EHDV isolates. The evolutionary relationship of GIV, KEMV, LIPV and TRBV with 14 other Orbivirus species was investigated, based on the ‘T2’ protein sequence. GIV VP2(T2) showed 23.7–56.2 % amino acid identity to the T2 of previously characterized orbiviruses, confirming its status as a member of a distinct Orbivirus species.

Previous serological studies have grouped members of the current Great Island virus species into two separate serocomplexes (Gorman et al., 1983). These are the Kemerovo serocomplex, which includes KEMV, LIPV and TRBV (tick and human isolates), and the Great Island serocomplex, which includes other tick- and seabird-derived isolates. This grouping agrees with the phylogenetic relationships between these viruses, as reported here.

The parameters used in the ‘polythetic definition’ of virus species within the family Reoviridae include: the ability of viruses to exchange (reassort) genome segments; serological relationships; genome segment migration patterns during agarose gel electrophoresis; similar sequence comparisons or cross-hybridization of RNA sequences; terminal conserved sequences on their different genome segments; and host range, vectors and the nature of the clinical signs induced (Mertens et al., 2005). Previous laboratory analyses have shown that reassortment frequencies between members within the Great Island serocomplex are high, whilst those between members of the Great Island and Kemerovo serocomplexes are significantly lower (Nuttall & Moss, 1989). Although further comparisons using sequence data for additional virus isolates would be important, consideration could be given to the reclassification of members of the Great Island virus species as two distinct but related species. The seabird/tick isolates (including GIV and BRDV) within the GIV serocomplex would remain within the species, whilst KEMV, LIPV and TRBV within the Kemerovo virus serocomplex could become members of a separate ‘Kemerovo virus’ species. It is notable that the highest level of amino acid identity previously detected between the T2 proteins of different Orbivirus species was between BTV and EHDV, two groups/species of virus that also show low levels of serological cross-reaction and evidence of intergroup reassortment during their evolution.

It has been suggested previously that the non-vectored dsRNA viruses have evolved by co-speciation with their respective hosts (Attoui et al., 2002). Within the T2 protein phylogenetic tree, SCRV appears to represent the oldest orbivirus lineage sequenced to date (Attoui et al., 2001), providing a ‘root’ for all of the other orbiviruses. So far, SCRV does not have any known vertebrate host and could be regarded as a true ‘tick virus’ (which infects ticks but is not transmitted to a vertebrate host, which is necessary for it to be considered an arbovirus), as per the definition provided by Nuttall (2009). Previous phylogenetic analyses, based on mitochondrial genes, have indicated that ticks also represent a root for other arthropods, including flies and mosquitoes (Wilson et al., 2000). A neighbour-joining analysis of orbivirus T2 proteins using a P-distance algorithm showed that a group including GIV, KEMV, LIPV, TRBV and BRDV (which are also tick-borne) provided a ‘root’ for the mosquito-borne orbiviruses. A similar situation was presented by the phylogenetic tree comparing the sequences of the orbivirus polymerase, where the tick-borne viruses provided a root for the insect-borne viruses. These findings suggest co-speciation between these viruses and their arthropod vectors.
Previous evolutionary analyses of ticks and insects have suggested that ticks appeared approximately 225 million years ago (MYA) (Klompen et al., 1996), whilst the earliest dating of culicine mosquitoes is about 150 MYA (Calvo et al., 2009). *Culicoides* biting midges are vectors for several orbiviruses and have been dated to the Cretaceous period (140–65 MYA) (Grimaldi & Engel, 2005; Grogan & Szadziewski, 1988). Considering the descending order for these dates and the topologies of phylogenetic trees for the orbivirus Pol and T2 proteins, a co-speciation hypothesis suggests that the earliest orbiviral ancestor existed at least 225 MYA and is likely to be a tick virus or tick-borne virus. It appears likely that a mosquito or mosquito-borne orbiviral ancestor evolved from this ancestral virus and that the *Culicoides* or *Culicoides*-borne viruses were the last to appear. This is similar to what has been suggested for the non-vectored aquareoviruses and orthoreoviruses, which seem to have co-speciated with their mammalian, avian, reptilian (orthoreoviruses) and aquatic (aquareoviruses) hosts and have evolved from a common ancestor that would have existed 520 MYA (Attoui et al., 2002).

**METHODS**

**Virus propagation, electron microscopy and purification of nucleic acids.** GIV (passage 5 in suckling mice, lyophilized in 1984), KEMV (passage 2 in Vero cells, lyophilized in 1980), LIPV (passage 3 in suckling mice, lyophilized in 1985) and TRBV (passage 16 in suckling mice, lyophilized in 1985) were grown in BHK-21 cells (clone BSR) at 37 °C in Glasgow minimum medium supplemented with 10% FBS, 100 IU penicillin ml⁻¹ and 100 μg streptomycin ml⁻¹. Infected cell cultures were incubated at 37 °C for 72 h until cell lysis began. The cells were then scraped into the supernatant and centrifuged at 3000 u for 10 min. The cell pellet was suspended in deionized water (to cause cell lysis) as described previously (Attoui et al., 2006). GIV was purified from the cell lysates using a discontinuous caesium chloride gradient, as described previously (Attoui et al., 2006). The purified virus was used for dsRNA extraction using RNA NOW reagent (Biogentex), as described previously (Attoui et al., 2000a, 2006). For electron microscopy, the caesium chloride-purified virus was dialysed in 10 mM Tris/HCl (pH 8.0) containing 150 mM NaCl and adsorbed to Formvar/carbon-coated grids. The grids were stained with 2% potassium phosphotungstate for 30 s, dried and then examined by transmission electron microscopy.

**Cloning of dsRNA segments.** GIV, KEMV, LIPV and TRBV genome segments were copied into cDNA, cloned and sequenced using a single-primer amplification technique as reported previously (Attoui et al., 2000a, b).

**Sequence comparison and analyses.** The sequence of the GIV genome was compared to a database containing the sequences available for different reoviruses, using the DNATools package (version 5.2.0.18; S. W. Rasmussen, Valby Data Center, Denmark). Sequence alignments were performed using CLUSTAL_X software (Thompson et al., 1994). Phylogenetic analyses were performed using MEGA4 software (Kumar et al., 2004) with the P-distance or the Poisson correction models and neighbour-joining method (Saitou & Nei, 1987) for tree building. The shape parameter x for the T2 sequence set was calculated using the PAML package (Yang, 1997).

Amino acid sequence comparisons with representative viruses from 14 genera of the family *Reoviridae* identified VP1 as the GIV Pol. The accession numbers used in these analyses are provided in Supplementary Table S3 (available in JGV Online). The 'T2 protein' of GIV was compared with T2 sequences from 12 *Orbivirus* species retrieved from the international sequence databases. Sequence accession numbers used in these analyses are provided in Supplementary Table S4.

**ACKNOWLEDGEMENTS**

This study was supported by funding from the Biotechnology and Biological Sciences Research Council and from Pfizer.

**REFERENCES**


Csorba, T., Bovi, A., Dalmay, T. & Burgyan, J. (2007). The p122 subunit of tobacco mosaic virus replicase is a potent silencing...


