Inter-segment complementarity in orbiviruses: a driver for co-ordinated genome packaging in the Reoviridae?

Mark Boyce,1 Malcom A. McCrae,1 Paul Boyce2 and Jan T. Kim1

1The Pirbright Institute, Pirbright, Woking GU24 0NF, UK
2Mott MacDonald, Mott MacDonald House, 8–10 Sydenham Road, Croydon, CR0 2EE

The process by which eukaryotic viruses with segmented genomes select a complete set of genome segments for packaging into progeny virus particles is not understood. In this study a model based on the association of genome segments through specific RNA–RNA interactions driven by base pairing was formalized and tested in the Orbivirus genus of the Reoviridae family. A strategy combining screening of the genomic sequences for inter-segment complementarity with direct functional testing of inter-segment RNA–RNA interactions using reverse genetics is described in the type species of the Orbivirus genus, Bluetongue virus (BTV). Two examples, involving four of the ten BTV genomic segments, of specific inter-segment interaction motifs whose maintenance is essential for the generation of infectious virus, were identified. Equivalent inter-segment complementarities were found between the identified regions of the orthologous genome segments of all orbiviruses, including phylogenetically distant species. Specific interaction of the participating RNA segments was confirmed in vitro using electrophoretic mobility shift assays, with the interactions inhibited using oligonucleotides complementary to the interaction motif of one of the interacting partners, and also through mutagenesis of the motifs. In each example, the base pairing rather than the absolute sequence was critical to the formation of a functional inter-segment interaction, with mutations only being tolerated in rescued virus if compensating changes were made in the interacting partner to restore uninterrupted base pairing. The absolute sequence of the complementarity motifs varied between species, indicating that this newly identified phenomenon may contribute to the observed lack of reassortment between Orbivirus species.

INTRODUCTION

The packaging of viral genomes requires the selective recognition of the viral genome amongst cellular nucleic acids. In the case of viruses that divide their genome between multiple separate molecules the additional requirement that a full set of genome segments is included in the virus particle should be satisfied to maximize the fraction of particles that are infectious. An understanding of how these discriminatory selections are achieved in virus families which have highly segmented genomes such as the Reoviridae is lacking.

The family Reoviridae consists of icosahedrally symmetrical viruses which have a dsRNA genome divided into 9 to 12 physically separate linear segments, the precise number depending on the genus. Following entry into the host cell the dsRNA segments remain enclosed within a partially disassembled capsid and are repeatedly transcribed by the viral transcription complexes (Martin & Zweerink, 1972; Verwoerd & Huismans, 1972; Verwoerd et al., 1972). The newly synthesized plus-sense transcripts are extruded into the cytosol where they function as mRNAs in translation of the viral proteins, and as replication intermediates which are packaged and replicated during virion assembly to form the progeny dsRNA genome (Patel & Roy, 2014; Zweerink et al., 1972). A single ORF is normally present on the sense strand, flanked by untranslated RNA and positioned between sequences of 5–10 nt length at each terminus that are conserved across all the genome segments (Joklik, 1983).

All of the virus genes in well-characterized genera of the Reoviridae are either essential or highly beneficial to the replication of the virus in its natural host, indicating that the packaging of a complete set of genome segments into a single virion is crucial to the generation of a viable virus. Members of the Reoviridae have a protein capsid which precisely defines the volume in which the genome is contained.
From this rigid architecture and the size of the genome it can be concluded that there is little additional space for packaging more than one copy of each segment. A consequence of this is that the strategy of packaging a number of molecules which substantially exceeds the total number of distinct segments to achieve at least one copy of each with high probability is not a possibility (Gouet et al., 1999). Under conditions of transmission at high m.o.i. the trans-complementation of two or more incomplete genomes can occur during the co-infection of a single cell, as suggested for influenza A viruses (Brooke et al., 2013). However, during transmission events at low m.o.i. the selective pressure for a mechanism which increases the probability of recruiting all genome segments is retained. Together, these considerations indicate that the members of the Reoviridae use a highly regulated mechanism which allows the specific recruitment of a single copy of each segment with high efficiency. Since the original characterization of the segmented nature of the genome it has been recognized that either protein–RNA interactions and/or RNA–RNA interactions underlie the mechanism for this segment-specific recruitment during assembly, but the nature of the molecular interactions enabling this precision are not known (Joklik, 1983).

The only segmented genome virus family for which there is an experimentally verified model of the key interactions driving genome packaging is the Cystoviridae family of three-segment bacteriophages. In this system each segment is specifically recognized by the pro-capsid which adopts two intermediate conformational states that successively generate specificity for the next segment to be packaged, in an ordered sequence (Mindich, 1999; Qiao et al., 1995). For the more highly segmented Reoviridae, a model similarly based on specific RNA–protein interactions may be considered unlikely as it would require up to 11 distinct intermediate binding specificities to be generated in the capsid. There is no evidence that such structurally distinct capsid variants exist amongst the known capsid structures of this family (Grimes et al., 1998; Lawton et al., 1997; Mindich, 1999; Nakagawa et al., 2003; Pesavento et al., 2006; Reinisch et al., 2000). Additionally, the highly reiterative nature of the structural proteins which define the architecture of the virus particle argues that they would be poor candidates for providing a unique copy of a binding site for each segment during genome packaging. By contrast, a model in which reconstitution of the genome is driven by sequence specific inter-segment RNA–RNA interactions is more attractive, as only one copy of each interacting RNA is required. Such a model predicts the existence of specific inter-segment sequence complementarities necessary for driving the association of the genome segments and therefore essential for the production of viable progeny virions. Against this background, the aim of this study was to develop a strategy combining bioinformatics-based and experimental approaches to begin the identification of biologically essential inter-segment interactions in the Orbivirus genus of the Reoviridae. Starting with the bioinformatics-based identification of all possible complementarities between genome segments of a defined length, a series of filters predicated on known biological characteristics of this genus were used to eliminate >99.99% of these. The remaining candidates were tested directly for their biological function using the reverse genetics system of the type species, Bluetongue virus (BTV), and were also characterized in vitro using electrophoretic mobility shift assays (EMSAs) (Boyce et al., 2008).

RESULTS

Bioinformatics-based analysis of all possible base-pairing interactions between segments

The generation of all possible base pairings between segments was performed by adapting RNAduplex of the ViennaRNA package 2 to scan along the genome segments in 1 nt steps identifying the most energetically stable potential interactions (Lorenz et al., 2011; Turner & Mathews, 2010). To limit less energetically favourable possibilities, such as those rich in A–U base pairs, non-canonical G–U base pairs and those with mispairing, a threshold in the predicted minimum free energy value was set at −7.0 kcal mol⁻¹. All possible 6 nt sequences in the genome of the South African reference strain of BTV type 1 were used as search terms in the sliding window scan of the entire genome. This returned 61 002 potential interactions with predicted stability greater than the threshold (Table 1).

Identification of potentially functional inter-segment complementarities

A series of filters based on the variants observed in the Orbivirus genus and related genera were applied successively to reduce the number of potential interactions to a number that could be tested functionally. First, based on analysis of defective interfering segments and other naturally occurring rearrangements of genome segments, members of the Reoviridae have been shown to consistently retain the 5′- and 3′-end regions in their normal position (Ballard et al., 1992; Feenstra et al., 2014; Gault et al., 2001; González et al., 1989; Gorziglia et al., 1989; Hua & Patton, 1994; Hendley et al., 1985; Matsui et al., 1990; Méndez et al., 1992; Pedley et al., 1984; Scott et al., 1989; Tian et al., 1993). This indicates that these regions contain all the cis-acting sequences needed to regulate packaging, genome replication and transcription of the segment. The importance of the terminal regions has been verified experimentally in the Orbivirus, Orthoreovirus and Rotavirus genera of the Reoviridae using reverse genetics, with terminal regions as short as 96 nt identified as sufficient to permit the recovery of the segment in a replicating virus (Boyce & McCrae, 2015; Boyce et al., 2012; Burkhardt et al., 2014; Matsuo & Roy, 2009; Roner & Joklik, 2001; Roner & Roehr, 2006; Roner & Steele, 2007a, b; Roner et al., 2004; Troupin et al., 2011). Based on these results the generation of a set of possible base pairings between segments was repeated with both
the source of the 6 nt search terms and the target for the searches confined to the 150 nt at the ends of each segment. This filter reduced the number of potential interacting sequences to 1840 (Table 1).

In the genus *Orbivirus* the 'species' taxon has the functional meaning that two virus isolates are in the same species only if they can exchange genome segments, when infecting the same cell, producing reassortant progeny which have segments from both parental viruses. This reassortment of genome segments shows that the *cis*-acting sequences required for selecting a complete set of segments are functionally conserved across different isolates within a species. Therefore, if inter-segment RNA–RNA interactions are required for segment selection, then the interacting regions can be expected to be conserved features. The identification of such conserved regions was carried out by the alignment of each segment from available BTV isolates with CLUSTAL W (Thompson *et al.*, 1994). Using these alignments, a second filter was introduced to eliminate sequences which were not within a conserved region in both partners, leaving 59 potential 6 nt interactions (Table 1).

The selection of a complete set of genome segments for packaging is a fundamental part of the replication cycle which should be expected to be conserved between different *Orbivirus* species, and this was used as the basis for the last filter. It was determined whether specific inter-segment complementarities were present in the same segments at equivalent positions when all *Orbivirus* species were compared with the type species, BTV. To achieve this, the mean minimum distance between potentially equivalent complementarities was calculated across the *Orbivirus* species using the termini of the segments as fixed reference positions. If both segments in the pair were found to contain potentially interacting sequences at a similar position in all species then the sequence was scored as conserved across all species. This final filter left six identified pairs of 6 nt sequences which, due to their overlapping nature, represented two inter-segment RNA–RNA interactions (Table 1). Together, these successively applied filters reduced the number of regions of complementarity identified to an experimentally testable number, revealing the existence of complementarities between segment 5 (S5) and segment 10 (S10) and between segment 1 (S1) and segment 7 (S7) that occur in all 15 *Orbivirus* species (Table 2).

### Sequence analysis of complementarities conserved across all orbiviruses

The conserved complementarities identified in the bioinformatic analysis between S5 and S10 and between S1 and S7 were aligned to identify features conserved at the level of the primary sequence (Fig. 1). In both cases the complementarity represented a 5' end to 3' end interaction. The length of the inter-segment complementarity was 6–8 nt for the S5:S10 pairs and 6–11 nt for the S1:S7 pairs, and in all cases the complementarity consisted of G–C and A–U base pairs, with no examples of non-canonical G–U base pairs. Of the four identified motifs, those in S5, S10 and S7 were located within the UTRs, whereas the motif in S1 was within the 5' end of the VP1 ORF, positioned 3–11 nt downstream of the initiation codon, depending on the species. Partial sequence conservation was observed across the *Orbivirus* genus, with 11/15 species containing the sequence 5′-RCCA-3′ in the S5 motif and 13/15 containing 5′-GUG-3′ in the S1 motif.

To be available for intermolecular base pairing, the identified motifs would be expected to be located in regions which are available for intermolecular interactions. To assess whether the complementarity motifs may occur within identifiable secondary structure contexts the most favourable secondary structures of the complete segments were predicted by minimum free energy using Mfold (Zuker, 2003). In each case one complementarity motif of the pair was predicted to occupy an unpaired loop at the apex of an extended stem, consistent with availability for intermolecular interactions; the S10 and S7 motifs were predicted to be loops positioned at the end of stable stems in each species, and the complementary S5 and S1 motifs were predicted to occupy regions of low intramolecular base pairing (Figs 2 and 3 and S1–S4, available in the online Supplementary Material).

### S5:S10 and S1:S7 complementarities are essential for the viability of BTV

If the identified complementarities are functionally essential to the recruitment of genome segments for packaging then the maintenance of base pairing between the two partners would also be expected to be essential. To investigate this, the viability of BTV with either disrupted

---

**Table 1.** Application of biological filters to identify sites of potential inter-segment base pairing

<table>
<thead>
<tr>
<th>Biological filter</th>
<th>No. of potential interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filtering by minimum free energy threshold (−7.0 kcal mol⁻¹)*</td>
<td>61 002</td>
</tr>
<tr>
<td>Filtering for 150 nt terminal regions†</td>
<td>1840</td>
</tr>
<tr>
<td>Filtering for sequence conservation within BTV‡</td>
<td>59</td>
</tr>
<tr>
<td>Filtering for positional equivalence in other species§</td>
<td>6</td>
</tr>
</tbody>
</table>

*Total number determined across the entire BTV genome using RNA-duplex with a minimum free energy threshold of −7.0 kcal mol⁻¹.
†Number within the 150 nt terminal regions of BTV genome segments.
‡Number which occur within conserved regions of BTV genome segments.
§Number having a mean minimum distance value of <20 nt when comparing the position of complementarities in all *Orbivirus* species with the BTV species. Complementarities within a segment are not included.
complementarities or with complementarities achieved using alternative sequences was tested using reverse genetics. Mutant S5 or S10 ssRNAs were made which contained 1, 2 or 3 nt changes within the interaction motif (Fig. 4a). In each case the recovery of these mutant segments with the wild-type partner in replicating virus was not possible. However, when the equivalent S5 + S10 double mutants were made with restored base pairing, each double mutant was recovered (Fig. 4a). This demonstrated that it was the complementarity between these two

Table 2. BTV inter-segment complementary sequences identified using the biological filters

| First segment | Second segment | Position (nt) and sequence | Distance from terminus (nt) | Position (nt) and sequence | Distance from terminus (nt) | Minimum free energy (kcal mol⁻¹) | Mean minimum distance (nt) |† |
|---------------|---------------|----------------------------|---------------------------|----------------------------|---------------------------|---------------------------------|---------------------------|
| S1            | S7            | 26–31 CGTGCA               | 26                        | 1075–1080 TGCAGG           | 76                        | −9.5                            | 7.15                       |
| S5            | S10           | 23–28 CAACCA              | 23                        | 754–759 TGTTGT             | 63                        | 8.2                             | 11.08                      |
| S5            | S10           | 24–29 ACCCAC             | 24                        | 753–758 GTGGTT             | 64                        | 7.4                             | 11.54                      |
| S5            | S10           | 25–30 ACCACC             | 25                        | 752–757 AGGTGT             | 65                        | −7.3                            | 12.00                      |
| S7            | S1            | 1075–1080 TGCAGG           | 76                        | 26–31 CCGTGCA             | 26                        | −8.6                            | 13.08                      |
| S7            | S1            | 1076–1081 GCAGGG           | 75                        | 25–30 CGTGCA             | 25                        | −10.6                           | 13.08                      |

•Position within the segment, sequence and distance from the nearest terminus given for the BTV species.

†Mean minimum distance when comparing the position of complementarities in all *Orbivirus* species with the position in the BTV species.

**Fig. 1.** Alignment of 15 *Orbivirus* species showing the identified inter-segment complementarity between the termini of (a) S5 : S10 and (b) S1 : S7. Dark shading, the extent of uninterrupted inter-segment complementarity (Watson–Crick base pairing); light shading, translation initiation codon; boxing, partial sequence homology in the complementarity motifs between virus species. Segment naming is harmonised to the type species (BTV) for consistency. Refer to Table 2 for the nucleotide positions of the complementarity motifs in the BTV species. PHSV, Peruvian horse sickness virus; EHDV1, epizootic hemorrhagic disease virus type 1; AHSV1, African horse sickness virus type 1; EEV, equine encephalitis virus; SCRV, St Croix River virus; GIV, Great Island virus.
motifs that was critical, rather than the absolute sequence. Similarly, mutants of S1 or S7 which were predicted to disrupt the interaction with wild-type S7 or S1 were only recovered as replicating virus when the compensating mutations were made to restore base pairing between the two partners (Fig. 4b).

**Fig. 2.** Minimum free energy secondary structure predictions of inter-segment complementarity motifs of the (a) S5 and (b) S10 segments of five Orbivirus species predicted using Mfold. The sequences of the complete genome segments were used to predict the minimum free energy arrangement of internal base pairing. Left to right: BTV type 1, epizootic hemorrhagic disease virus type 1, African horse sickness virus type 1, St Croix River virus and Peruvian horse sickness virus. The minimum free energy associated with these base-paired regions is indicated in kcal mol$^{-1}$. Black bases, extent of complementarity between S5 and S10 segments. Numbering indicates the position within the genome segment. Secondary structure predictions of the remaining virus species are shown in Figs S1 and S2.
**S5:S10 and S1:S7 interactions occur in the absence of other viral factors**

The absence of understanding of the temporal organization of RNA–RNA and protein–RNA interactions involved in genome packaging in the Reoviridae dictates that the stage of genome segment selection that the identified interactions represent is unknown. However, to investigate whether the two identified interactions could occur in the

---

**Fig. 3.** Minimum free energy secondary structure predictions of inter-segment complementarity motifs of the (a) S1 and (b) S7 segments of five Orbivirus species predicted using Mfold. The sequences of the complete genome segments were used to predict the minimum free energy arrangement of internal base pairing. Left to right: BTV type 1, epizootic hemorrhagic disease virus type 1, African horse sickness virus type 1, St Croix River virus and Peruvian horse sickness virus. Labelling as described in Fig. 2. Secondary structure predictions of the remaining virus species are shown in Figs S3 and S4.
absence of the other viral ssRNAs and proteins, the potential for the two partners to associate in solution was tested by EMSA under native conditions. ssRNA transcripts were generated by run-off transcription from T7 promoter-driven clones as described previously (Boyce et al., 2008). When incubated together, the S5 and S10 ssRNAs were found to generate a slower migrating complex and visible depletion of the free RNAs (Fig. 5a). Similarly, the incubation of the S7 ssRNA with the S1 ssRNA produced a retarded complex and the visible depletion of the free RNAs (Fig. 5b). Purification of each of the slow-migrating complexes and the free RNAs and their analysis under denaturing conditions confirmed that the complexes were composed of the two co-incubated ssRNAs (Fig. S5). This maintenance of the interaction under the prolonged conditions of the EMSA demonstrated that the S5 and S10 ssRNAs and the S1 and S7 ssRNAs could stably associate in the absence of other viral RNAs or proteins. To test the relevance of the complementarity motifs identified by the sequence analyses, the disruption of these specific interactions was analysed in vitro using antisense RNA oligonucleotides targeted to the extended stem–loop motif of each interacting pair. In each case the formation of the complex was inhibited in a dose-dependent manner by the competing oligonucleotide (Fig. 5c, d), demonstrating that the identified stem–loop regions within each full-length

![Fig. 4](image)

**Fig. 4.** Inter-segment complementarity is required for the recovery of BTV. Virus-induced cytopathic effect visualized by crystal violet staining. Shading, the extent of the complementarity motif; underlined, mutated bases. (a) S5 mutants and S10 mutants tested with WT S10 or WT S5 and with the compensating complementary mutants of S10 or S5. (b) S1 mutants and S7 mutants tested with WT S7 or WT S1 and with the compensating complementary mutants of S7 or S1. mut, Mutant.
ssRNA were necessary to drive the interaction. To evaluate the effect of the point mutations analysed by reverse genetics, the WT segments were compared with the corresponding point mutants in their formation of a complex. In each case the combinations that were incompatible during recovery by reverse genetics produced less of the RNA complex, consistent with these bases being determinants of the interaction between the segments during genome segment selection (Fig. 5e–h).

**DISCUSSION**

The aim of this study was to develop a strategy that would allow the identification of essential base-pairing interactions between genome segments in the *Orbivirus* genus and to test them for functionality. The conserved base-pairing interactions between specific segments within the 19 kb genome of all *Orbivirus* species have been found by applying criteria based on the known biology of the orbiviruses to form potential RNA–RNA interactions which satisfy a minimum free energy threshold. Testing these potential interactions using both the reverse genetics system (Boyce et al., 2008) and also the formation of stable complexes in vitro showed them to be biologically essential.

The combined bioinformatics and experimental approach adopted has defined two examples, involving four of the 10 genome segments, of inter-segment interactions which are conserved in all member species of the *Orbivirus* genus. In the type species BTV, the complementarities between these motifs were found to be essential for completion of the virus replication cycle. The reverse genetics approach used to test the requirement for sequence complementarity in these inter-segment RNA–RNA interactions demonstrates that the complementarities are biologically relevant. It remains a formal possibility that the essential inter-segment interactions, identified for the first time in this study, are required for those steps in the replication cycle outside of genome segment selection in which the ssRNA transcripts participate, i.e. translation and negative strand synthesis. However, the existence of the identified complementarities at the apex of predicted stem–loop structures provides a mechanism for how the segmented genomes may be reconstituted, whereas the other steps in which the ssRNAs participate in the replication of the *Reoviridae* do not predict the existence of such complementarities.

The complementarity motifs identified in this study share a number of features in both their primary sequence and predicted secondary structure. The positions of the motifs relative to the end of the segment are highly conserved; when considering all 15 virus species, the S1 motifs occur within 7 nt of each other, the S7 motifs within 13 nt, the S5 motifs within 9 nt and the S10 motifs within 39 nt (Fig. 1). Examination of the minimum free energy secondary structures shows similar predicted secondary structure contexts, with the S7 and S10 motifs occupying unpaired loops positioned at the end of long stable stems, in agreement with being available to interact with a partner sequence (Figs 2 and 3). The partial sequence conservation found across the *Orbivirus* genus in all four complementarity motifs, but absent from the adjacent sequences, is consistent with them being under greater selective pressure to remain unchanged than the flanking sequence (Fig. 1). This conservation suggests that these motifs have altered less rapidly during the radiation of the *Orbivirus* genus and conforms to the prediction that the simultaneous mutation of both interacting partners is needed to maintain the base pairing required for their function. Using reverse genetics, the alteration of these motifs demonstrated that single nucleotide mismatches within both the S5:S10 and S1:S7 pairs were sufficient to prevent the recovery of infectious virus, whilst more extensive changes in the sequence of the complementarity motifs were possible if the base pairing was maintained by simultaneously introducing a compensating change in the partner motif (Fig. 4). This is consistent with recent scanning mutagenesis of the terminal regions of BTV S10, which identified the interacting loop as a sequence which could not be varied without lethality, being positioned at the end of a stem which had little constraint on its absolute sequence composition, beyond being predicted to form a stem (Boyce & McCrae, 2015). The complete absence of non-canonical G–U base pairs in the 30 defined interactions (Fig. 1), in combination with the mutagenesis data (Fig. 4), demonstrates that the stability of the interaction between the motifs is important for their function. Therefore, RNA–RNA incompatibilities can be added to the already known protein–protein incompatibilities and protein–RNA incompatibilities which are believed to prevent the exchange of genome segments between different *Orbivirus* species, supporting a further functional underpinning for speciation in the genus (Grimes et al., 1998; Lawton et al., 1997; Modrof et al., 2005; Reinisch et al., 2000).

Virion assembly in the *Reoviridae* is initiated within electron-dense structures termed ‘viral inclusion bodies’ (VIBs) or ‘viroplasms’ which consist of a matrix of one or more non-structural proteins that recruits both structural proteins and the viral ssRNAs (Joklik, 1983). However, there is very little understanding of the processes involved in genome segment selection and packaging. After the synthesis and release of nascent viral ssRNAs by the infecting core particle, the next characterized stage is the appearance of newly assembled particles already consisting of the inner layer of the capsid, replicated (double-stranded) RNA segments and the viral RNA-dependent RNA polymerase/capping enzyme complex, exemplified by early biochemical studies of the *Orthoreovirus* genus (Morgan & Zweerink, 1975; Zweerink, 1974; Zweerink et al., 1972). Between these two stages, several processes must have occurred in order to select and package a viable genome. These include (1) selection of a complete set of ssRNA segments, (2) association of the complete set of ssRNAs with
Inter-segment complementarity in orbiviruses

(a) S5–S10 complex

(b) S1–S7 complex

(c) S5–S10 complex

(d) S1–S7 complex

(e) S5–S10 complex

(f) S1–S7 complex

(g) S5–S10 complex

(h) S1–S7 complex

Inter-segment complementarity in orbiviruses
structural proteins during virus assembly and (3) synthesis of the minus-sense strands to generate the dsRNA genome. The essential inter-segment interactions described here are consistent with a model in which sequence-specific interactions drive the association of the genome segments in their single-stranded form. Considering what models of specific RNA–RNA interaction might be operating, simple models based on the head-to-tail interaction of the segments are consistent with the data, as are more complex arrangements of the participating RNAs where some segments make multiple contacts consisting of short complementarities with other segments, rather than fewer more extensive contacts of the type characterized here. The involvement of conformational changes initiated by the interaction between segments and the potential of such changes to create new specificity for interacting with further segments are also a possibility, and would allow an order to be imposed on the process. In addition to the specific RNA–RNA interactions described, there must be interactions between the RNA and the protein components of the assembling virion; although not understood at the molecular level, these have been proposed to occur in orbiviruses in a sequential manner, similar to that observed in the Cystoviridae family of three-segment bacteriophages (Mindich, 1999; Qiao et al., 1995; Sung & Roy, 2014).

In any RNA–RNA interaction-driven model for specific genome segment selection, at least a further seven specific inter-segment interactions would be required to allow all 10 segments to interact, therefore it is instructive to consider why the screening filters applied in this first analysis did not identify other interacting segments, beyond the four analysed in detail. The two pairs of interaction motifs experimentally verified in this study were both continuous runs of at least seven bases where disruption of a single base pair was lethal for virus viability. This suggests that the base pairing is of a length that is sufficiently stable to initiate the interaction, but not sufficient to lock the RNAs in a complex which would preclude them successfully acting as mRNAs during translation or as templates during negative strand synthesis. Such a degree of stability could alternatively be achieved through discontinuous interaction motifs or multiple short motifs as well as the longer continuous examples identified in this study, in which case more sophisticated screening algorithms will be required to identify these shorter motifs. Additionally, the criteria used at each step of the bioinformatic screening process (Table 1) can be varied to analyze viral genome datasets such that more candidate interactions are generated. Although a rise in false-positives is an expected consequence of relaxing the criteria, variation of the criteria will be explored to determine whether further continuous motifs are identified.

The findings of this study show that the direct interaction between different viral segments in the ssRNA form is a crucial step in the assembly of viable virus within the VIBs. Short single-stranded loops are established as the initiators of specific interactions between RNA molecules in many biological systems, such as the anticodon stem–loop of tRNA molecules interacting with mRNA codons in the context of the ribosomal decoding centre (Crick, 1966). The location of the orbivirus complementarity motifs at the apex of predicted stem–loops on either one or both interacting partners is consistent with a model where interaction between RNAs is nucleated via these base-pairing interactions. The formation of a multi-segment complex suggests that these short duplex regions persist and are possibly stabilized either through additional RNA–RNA interactions or through interaction with a viral protein in the VIB. In both the S5:S10 and S1:S7 pairs the sequence permits the formation of additional energetically favourable base pairing adjacent to the complementarity motifs (Fig. S6), raising the possibility that the base pairing extends beyond the nucleation region of the described complementarity motifs.

The complementarity motifs described in this study are the first examples of inter-segment interactions which are essential for virus replication in segmented genome viruses. An inter-segment RNA–RNA interaction has recently been identified in a second virus family whose members have highly segmented genomes, the Orthomyxoviridae family. In that case extensive disruption of the interaction by
mutagenesis reduced the infectious titre of the virus by \(\sim 1 \log_{10}\), in contrast to the lethal phenotype observed in BTV in this study (Gavazzi et al., 2013). The interaction was found between two segments in closely related isolates of influenza A viruses, suggesting a role in a subset of related isolates which are able to re assort segments more freely with each other than with other isolates of the influenza A viruses species which lack the sequence complementarity (Essere et al., 2013; Gavazzi et al., 2013).

The essential interactions between genome segments reported in this study and the underlying complementarity motifs are a newly described feature of the replication cycle of the Reoviridae family, and represent a potential target for antiviral drugs. The formation of these interactions is critical for the viability of the virus, as shown by the lethality of single base mismatches (Fig. 4) and supported by the absence of G–U base pairs in the 30 examples encompassed by the 15 species. Taken together, the data strongly suggest that specific disruption of an inter-segment interaction would severely reduce/abrogate virus replication. The identification of a potential drug target in the Reoviridae family has particular relevance to the human mortality caused by rotaviruses which are responsible for \(\sim 500,000\) infant deaths per year worldwide (Parashar et al., 2006) and also the impact on high-value animals caused by some orbiviruses, such as African horse sickness virus which has a mortality rate of \(\sim 90\%\) in horses (Mellor & Hamblin, 2004).

The observations that viability is maintained when the interactions are achieved using alternative complementary sequences and that single-nucleotide mismatches are lethal suggests a way to design attenuated vaccines in this virus family which do not exchange genome segments with field strains. With inter-segment interaction achieved through alternative sequences with multiple changes in both motifs, such as those reported in this study (Fig. 4), the exchange of the engineered segments with circulating strains would be inhibited due to the incompatibility of the complementarity motifs. If extended to include all genome segments, such non-reassorting viruses would be genetically isolated from field strains, preventing the unintended introduction of vaccine-derived segments into the pool of circulating viruses during a vaccination programme.

In the wider context of the increasing appreciation of the roles that intermolecular RNA–RNA interactions have in biological processes, it may be expected that other segmented virus families use RNA–RNA interactions as an essential step in the regulated recruitment of their complement of genome segments. The combined approach of using biologically informed stepwise filtering of energetically favourable potential interactions and experimental verification could be applied more generally to identify such functionally essential inter-segment interactions in segmented viruses.

**METHODS**

**Viral sequences.** The sequences of BTV genome segments were downloaded from the National Center for Biotechnology Information (NCBI) Nucleotide database in November 2013. Incomplete segments and sequences containing ambiguities were removed from the dataset, leaving 69–162 examples of each genome segment.

The complete genomes of 15 Orbivirus species present in the NCBI Nucleotide database were downloaded in November 2013. The species of the Orbivirus genus were: African horse sickness virus, Bluetongue virus, Changuinola virus, Chuzan virus, Corripiorata virus, Epizootic hemorrhagic disease virus, Equine encephalitis virus, Great Island virus, Pata virus, Peruvian horse sickness virus, St Croix River virus, Tilligerry virus, Umatilla virus, Wallal virus and Yunnan virus. The accession numbers of the genome segments are listed in the supplementary material. Wallal virus sequences lacked \(\sim 10\) nt at the termini of each segment and therefore were omitted from the calculation of the positional conservation of complementary motifs.

**Identification of potential RNA–RNA interactions in Orbivirus species.** A script was developed to extract the hexanucleotide at each position along the entire virus genome and to use the RNAduplex program of the Vienna RNA package (Lorenz et al., 2011) for computing a duplex between each hexanucleotide and all segments of the genome to find the energetically optimal inter-segment hybrids, thereby identifying regions of complementarity of \(\geq 6\) nt.

**Measuring sequence conservation between BTV isolates.** Multiple sequence alignments were constructed for each genome segment using CLUSTAL W (Thompson et al., 1994). The \(6\) nt words were scored to identify those that were highly conserved by calculating the arithmetic mean of the Shannon entropies at the alignment columns they spanned (Capra & Singh, 2007; Valdar, 2002). A mean entropy threshold of 0.01 bits was used to define conservation of potential interacting sequences, on a scale where 0 corresponds to perfect conservation and 2 indicates maximal randomness.

**Measuring positional conservation of potential interacting sequences across Orbivirus species.** Given two potential inter-actions between positions \(x_1\) and \(x_2\) in species \(x\) and positions \(y_1\) and \(y_2\) in species \(y\), the distance is \(|x_1–y_1|+|x_2–y_2|\), if \(x_1\) is in the same end of the same segment as \(y_1\) and the same is true for \(x_2\) and \(y_2\). On this basis, for all potential interactions found in the BTV genome, a closest potential interaction in each of the other Orbivirus species was determined and the arithmetic mean of these minimal distances (the ‘mean minimum distance’) was calculated. The source code for this software is available on request.

**Prediction of intramolecular secondary structures.** The prediction of intramolecular base pairing by minimum free energy was performed with complete genome segments using Mfold (Zuker, 2003). Wallal virus sequences lacked \(\sim 10\) nt at the termini of each segment and therefore were omitted. The prediction of the minimum free energy associated with specific stems was performed with RNAeval of the Vienna RNA package (Lorenz et al., 2011).

**Cell lines and virus.** BSR cells (a BHK-21 subclone) were cultured in Dulbecco’s modified Eagle’s medium supplemented with 5% (v/v) FBS at 36°C in 5% CO\(_2\), and were used in the generation of WT and mutant BTVs by reverse genetics.

**Synthesis of viral ssRNAs.** Synthetic ssRNAs were prepared by run-off in vitro transcription using T7 RNA polymerase either from plasmid clones which had been linearized with BsmBI, BsaI or BpiI, or from PCR products generated using KOD DNA polymerase (Merck Millipore). Transcripts were prepared with anti-reverse cap analogue
(ARCA) from the linear DNA using a mMESSAGE mMACHINE T7 Ultra kit (Ambion) as described previously (Boyce et al., 2008). Antisense RNA oligonucleotides were prepared by run-off in vitro transcription using T7 RNA polymerase from PCR products, without ARCA.

**EMSA.** Pairs of purified ssRNAs (0.35 pmol each) were incubated at 20 °C in 6 μl PBS + 2 mM MgCl₂ for 45 min. Samples were separated on 1.0 % agarose gels containing 0.005 % (w/v) ethidium bromide after addition of 1.5 μl loading buffer [50 % (v/v) glycerol, 0.05 % (w/v) xylene-cyanol FF]. Native gel electrophoresis of the RNA complexes was performed at 4 °C in a buffer containing 40 mM Tris, 20 mM acetic acid and 0.25 mM MgCl₂ (pH 8.4).

In competition experiments, the antisense RNA oligonucleotide was pre-incubated with the target ssRNA for 20 min prior to incubation with the partner ssRNA for 45 min.

**Purification of RNA complexes and analysis by denaturing agarose gel electrophoresis.** The complexes were excised from native gels and purified using RNaid (MP Biomedicals). The purified complexes were analysed under denaturing conditions by electrophoresis on 1 % agarose in MOPS electrophoresis buffer in the presence of formaldehyde (Sambrook & Russell, 2001).

**Transfection of BSR cells to generate mutant BTVs by reverse genetics.** The variants of S1, S5, S7 and S10 were rescued in the BTV type 1 South African reference strain genetic background using the BSR cell line, as described previously (Boyce et al., 2008).

**Visualization of viral cytopathic effect.** At 5 days post-infection monolayers were washed in Hank’s balanced salt solution and fixed with 10 % formaldehyde in PBS for 30 min. Staining was performed with 0.2 % crystal violet, 20 % ethanol for 2 min. Stained monolayers were rinsed three times in water and dried.

**ACKNOWLEDGMENTS**

We thank Paul Boyce for his expertise in developing algorithmic approaches to searching for inter-segment complementarity.

**REFERENCES**


