Rapid mapping of functional cis-acting RNA elements by recovery of virus from a degenerate RNA population: application to genome segment 10 of bluetongue virus

M. Boyce and M. A. McCrae

The Pirbright Institute, Woking GU24 0NF, UK

The regulatory elements which control the processes of virus replication and gene expression in the Orbivirus genus are uncharacterized in terms of both their locations within genome segments and their specific functions. The reverse genetics system for the type species, Bluetongue virus, has been used in combination with RNA secondary structure prediction to identify and map the positions of cis-acting regions within genome segment 10. Through the simultaneous introduction of variability at multiple nucleotide positions in the rescue RNA population, the functional contribution of these positions was used to map regions containing cis-acting elements essential for virus viability. Nucleotides that were individually lethal when varied mapped within a region of predicted secondary structure involving base pairing between the 5’ and 3’ ends of the transcript. An extended region of predicted perfect base pairing located within the 3’ untranslated region of the genome segment was also found to be required for virus viability. In contrast to the identification of individually lethal mutations, gross alteration of the composition of this predicted stem region was possible, providing the base-pairing potential between the two strands was maintained, identifying a structural feature predicted to be conserved throughout the Orbivirus genus. The approach of identifying cis-acting sequences through sequencing the recovered virus following the rescue of a degenerate RNA population is broadly applicable to viruses where reverse genetics is available.

INTRODUCTION

The replication and expression of viral genomes is regulated by a functionally diverse range of cis-acting elements. For segmented genome viruses which divide their genome between multiple separate molecules there is the additional requirement that each genome segment has its own cis-acting elements to regulate these functions.

The family Reoviridae consists of icosahedrally symmetrical viruses which have a dsRNA genome divided into nine to 12 physically separate linear segments, the precise number being dependent 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 capped and extruded into the cytosol where they function both as mRNA in translation of the viral proteins and as replication intermediates which are packaged and used as templates for negative-strand synthesis during the assembly of progeny virions. Within the Reoviridae little is known about either the location or functioning of cis-acting sequence elements which regulate these and other virus-directed processes which must operate at the level of either the dsRNA genome segments or the single-stranded transcripts made from them. This is in contrast to the structure–function relationships of the viral proteins which have been extensively characterized in the genera Rotavirus, Orbivirus and Orthoreovirus (Desselberger, 2014; Nibert, 1998; Patel & Roy, 2014; Stencel-Baerenwald et al., 2014). In all but a small minority of genome segments, a single ORF is present on the sense strand, located between untranslated regions (UTRs) which vary in length and sequence between the segments (Joklik, 1983). The sequences regulating important viral processes such as transcription, genome replication, differential gene expression, RNA stability, genome assortment and genome packaging are yet to be identified or investigated at the mechanistic level.

Defective interfering segments and other naturally occurring rearrangements of genome segments in members of the Reoviridae have been shown to consistently retain the...
5’ and the 3’ end regions of the affected genome segment (Ballard et al., 1992; Desselberger, 1996; Feenstra et al., 2014; Gault et al., 2001; González et al., 1989; Gorziglia et al., 1989; Hua & Patton, 1994; Hundley et al., 1985; Matsui et al., 1990; Méndez et al., 1992; Palombo et al., 1998; Pedley et al., 1984; Scott et al., 1989; Shen et al., 1994; Tian et al., 1993). In each case the two regions are retained in their normal positions at the termini of the segment, indicating that they contain the cis-acting sequences which regulate transcription, genome replication and packaging of the segment. The importance of the terminal regions has been verified experimentally in the genera Orbivirus, Orthoreovirus and Rotavirus, by recovering mutant segments which incorporate a pre-defined rearrangement or deletion, using reverse genetics (Boyce et al., 2008). Using this system, the terminal 150 nt at each end of segment 10 have been shown to be sufficient to regulate packaging, genome replication, transcription and expression of a reporter gene inserted between these terminal regions (Boyce et al., 2012). To determine whether cis-acting elements occupy a contiguous stretch of sequence within the termini of genome segments or are distributed within the terminal regions, a scanning mutagenesis strategy was developed to provide the detailed resolution not achievable through the generation of deletion mutants. Here, an approach based on the reverse genetics system can be manipulated at the sequence level using the helper-independent reverse genetics system (Boyce et al., 2008). Using this system, the terminal 150 nt at each end of segment 10 have been shown to be sufficient to regulate packaging, genome replication, transcription and expression of a reporter gene inserted between these terminal regions (Boyce et al., 2012). To determine whether cis-acting elements occupy a contiguous stretch of sequence within the termini of genome segments or are distributed within the terminal regions, a scanning mutagenesis strategy was developed to provide the detailed resolution not achievable through the generation of deletion mutants.

The type species of the Orbivirus genus, Bluetongue virus, can be manipulated at the sequence level using the helper-independent reverse genetics system (Boyce et al., 2008). Using this system, the terminal 150 nt at each end of segment 10 have been shown to be sufficient to regulate packaging, genome replication, transcription and expression of a reporter gene inserted between these terminal regions (Boyce et al., 2012). To determine whether cis-acting elements occupy a contiguous stretch of sequence within the termini of genome segments or are distributed within the terminal regions, a scanning mutagenesis strategy was developed to provide the detailed resolution not achievable through the generation of deletion mutants. Here, an approach based on the reverse genetics system was used to simultaneously probe multiple sites within the terminal regions of genome segment 10. In order to discriminate between functionality at the level of the primary sequence and participation in potential secondary structure features, the identity of bases which are lethal when varied was analyzed in combination with the predicted secondary structure of the transcript. The data generated from such a combined strategy can be used both to inform subsequent studies of specific motifs and for the investigation of the molecular mechanisms by which they regulate viral processes.

RESULTS

Introduction of non-coding changes at defined intervals using degenerate primers

The approach used was to recover virus generated using one RNA transcript containing degeneracy at specific positions and the remaining nine wild-type (WT) transcripts. The systematic introduction of point mutations into the segment 10 transcript was performed by using degenerate PCR primers to produce a T7 transcription template with the form T7 promoter- segment 10, required for use in the bluetongue virus reverse genetics system (Boyce et al., 2008) (Fig. 1a). The presence of the degeneracy in the PCR product was assessed by sequencing the targeted region and confirming the presence of the variant bases at the degenerate positions (Fig. 1b). In the UTRs, the choice of positions for mutation was unrestricted, whereas within the ORF the redundancy of the genetic code was used to only introduce silent mutations. A binary choice between the WT base and the mutant base was used at each position such that the alternative base was selected to cause the maximum disruption to potential internal secondary structures formed through intramolecular base pairing.

Analysis of the 3’ end of segment 10 using primary scans

An initial primary scan of the 3’ end of segment 10 was performed at 12 nt intervals, to achieve a coarse map of the region from positions 685 to 757, with seven positions being simultaneously interrogated (positions 685, 697, 709, 721, 733, 745 and 757), using primer Degen 1a 3’. This region includes the carboxy terminus of the NS3 ORF, where silent alternative bases in terms of coding potential were used, and extends 50 nt downstream of the termination codon (Fig. 1a). The recovered virus population was titrated and individual plaques were picked and propagated. The region subjected to degenerate mutagenesis was then amplified by reverse transcription (RT)-PCR and the amplicons were sequenced. Recovery of the variant/WT base at each degenerate position in the interrogated region was scored across each sequenced virus (n=19). Positions 685, 697, 709 and 745 were found to tolerate the variant base, while at positions 721, 733 and 757 only the WT base was recovered in progeny viruses (Table 1). This demonstrated that the bases required for virus viability do not form a contiguous region within the 3’ end of segment 10, a distribution that would not be revealed using a conventional deletion mutagenesis approach.

To complete the primary scan of the 3’ end of segment 10, the remainder of the 3’ UTR was mapped at 12 nt intervals from position 751 to the extreme 3’ terminus at position 822, with seven positions (751, 763, 775, 787, 799, 811 and 822) being simultaneously interrogated using primer Degen 1b 3’. In this scanned region, position 763 tolerated the variant base, while positions 751, 775, 787, 799, 811 and 822 always tolerated the WT base in the sequenced progeny viruses (n=18) (Table 1).

Mapping the 3’ end of segment 10 at greater resolution

The two primary scans allowed the construction of a coarse map identifying bases which do not tolerate the introduced variation (Fig. 2a). To obtain a more detailed map and relate it to the secondary structure of segment 10, the minimum free energy configuration of intramolecular base pairing was predicted using Mfold (Zuker, 2003) (Fig. 2b), and the predicted structure used to guide the choice of which bases to vary in secondary scans. Positions 703, 715,
727, 739, 745, 751, 754, 760, 766, 771, 781, 793, 805 and 817 were interrogated using the two degenerate primers Degen 2a 3’ and Degen 2b 3’, generating a higher resolution map of invariant bases within the predicted secondary structure (Fig. 3). These mapping data revealed that all the examples of variable positions within nt 715–822 occurred within a predicted stem formed by base pairing between nt 736–752 and nt 761–776 of the 3’ UTR (Fig. 3).

Mapping of the 5’ end of segment 10

Using the same sequential approach of primary and secondary scans, the 5’ end of segment 10 was scanned without the introduction of protein-coding changes using primers Degen 1a 5’, Degen 1b 5’, Degen 2a 5’ and Degen 2b 5’. Positions 3, 9, 15, 25, 28, 31, 32, 34, 37, 40, 41, 43, 46, 49, 52, 55, 67, 73, 76, 77, 79, 88, 100, 112, 124 and 136 were interrogated, which includes the amino terminus of the NS3 ORF beginning at position 20 (Fig. 1a). Within the 5’ end, the distribution of the invariant bases clustered towards the extreme 5’ end, in the region predicted to be able to form intramolecular stems with invariant regions located in the 3’ end of the segment (Fig. 4). Together, the invariant regions from both ends of the RNA molecule are predicted to form a region of extended base pairing in the single-stranded form, outside which the positions tested are all variable (Fig. 4).

Targeted mutagenesis of the segment 10 stem

The predicted long stem consisting of nt 736–752 and nt 761–776 in the 3’ UTR has a counterpart within the 3’ end of segment 10 of all Orbivirus species (Fig. S1, available in the online Supplementary Material), suggesting that this predicted structural feature is functionally important across the genus. In bluetongue virus, the scanning mutagenesis showed that individual positions within the stem can be varied without lethality (positions 739, 745, 763 and 771) (Fig. 3). Together these results indicated that single nucleotide mismatches are not sufficient to disrupt the formation of this stem. To test the functional importance of this predicted secondary structure feature directly,
The simultaneous introduction of multiple mutations, each with the WT base as an alternative, allowed the determination of whether each position varied is an essential base. The two primary scans covering the region from nt 685 to the 3′ terminus (nt 822) of segment 10 showed that the 3′ end consists of at least three regions containing invariant bases interspersed with at least three regions containing bases which tolerate variation without lethality (Fig. 2). The extra resolution generated by including the data from the secondary scans across the 3′ end region confirmed this distribution and revealed additional detail indicating a tolerance of variation within the stem portion of the predicted stem–loop structure at nt 736–776 (Fig. 3). The loop at the apex of this long predicted stem did not tolerate variation, with positions 751, 754, 757 and 760 each returning only the WT base in all viruses sequenced. In contrast, positions 739, 745, 763 and 771 within the stem could be varied, demonstrating that the stem is able to tolerate single nucleotide mismatches without losing its function. The recovery of extensively altered variants of this predicted stem was also possible providing that compensating mutations were made in the opposite strand, such that a virus with 16 nucleotide changes which maintained the stem was viable (Fig. 5). The tolerance of this region to major sequence changes, providing that base pairing is retained, is consistent with it being a stable secondary structure feature, where the stem is the functional entity and its base composition is less important. This is in contrast to single nucleotide changes which each individually confer a lethal phenotype, such as all other positions interrogated between nt 715 and 822 using the degenerate mutagenesis approach (Fig. 4), showing that these positions either act at the sequence level or participate in secondary structures which are more sensitive to disruption. Discriminating between these two categories, (1) elements which are functionally destroyed by a single base change and (2) stable structural elements which tolerate some mismatches without loss of function, relied on combining the mapping approach with the prediction of potential stable intramolecular stems which may be functional, and their subsequent analysis by directed mutagenesis. The loop positioned at the apex of the stem did not tolerate single nucleotide substitutions and is highly conserved at the sequence level among all bluetongue virus isolates (data not shown), suggesting that the predicted loop is a functionally important element where the sequence is essential for its function.

Within the 5′ end of segment 10 there was a division in the distribution of the bases which tolerate variation and those which are functionally essential. Variable bases were located 3′ of position 34, whereas the functionally essential bases mapped at the 5′ end of the scanned region (Fig. 4). The secondary structure prediction of the potential intramolecular base pairing within segment 10 placed the functionally important regions of the 5′ end within stems formed through base pairing with functionally important regions located at the 3′ end of the genome segment.

### DISCUSSION

This study used the reverse genetics system of the type species of the Orbivirus genus, Bluetongue virus (Boyce et al., 2008), to introduce variation into a genome segment during virus rescue in order to map cis-acting bases in the segmented genome. The combined approach of using the mapping of lethal mutations with predicted secondary structures has allowed the identification of both essential regions which are sensitive to single nucleotide changes and an example of a feature which is predicted by minimum free energy considerations and targeted mutagenesis to be functional at the level of secondary structure.

The mutagenesis technique uses the lethality conferred by mutations within essential sequences, such as a necessary cis-acting regulatory element, to map those sequences.
In total, the combination of the mapping data and the secondary structure prediction identified an extended region of potential secondary structure which includes all the functionally important bases, nt 1–43 and nt 715–822, with the interrogated bases lying outside these regions all found to be non-essential (Fig. 4).

The availability of sequence information from multiple virus isolates in public databases allows the comparison of molecular epidemiological data with the mapping results obtained in this study. This reveals that for the 28 alternative mutant bases which were not recovered by scanning mutagenesis the alternative nucleotides were also not

**Fig. 2.** Primary scan of the 3’ end of segment 10. Distribution of the variable and invariable positions within the interrogated region of segment 10 (a) as a linear map, and (b) mapped onto the minimum free energy secondary structure prediction of the segment 10 transcript, generated using Mfold. Numbering denotes the nucleotide position within segment 10 and arrows indicate the interrogated positions. Green highlighting shows positions which tolerated variation and red highlighting shows positions which returned no examples of the variant base.
represented in the sequences deposited in GenBank. In two of these varied positions (nt 721 and nt 775) a variant base other than the alternative base used in the scanning mutagenesis was represented in the deposited sequences. However, in both these cases the variant base found in the database sequences was able to form an alternative base pair that was consistent with the predicted secondary structure (Fig. 4) (G at nt 721 and U at nt 775), whereas the base chosen for mutagenesis was selected for maximal disruption to potential secondary structures and was unable to form an alternative base pair, showing full agreement between these two approaches.

The mapping approach described is designed to locate any functionally essential region within the interrogated part of the segment, at a level of resolution determined during design of the degenerate oligonucleotides. Here it has been used as a proof-of-principle study to define the location of cis-acting RNA elements, by not altering the coding potential of the genome segment, but may also be applied to interrogating the amino acid variation tolerated at defined positions in a protein. Because the approach relies on the propagation of viable virus recovered by reverse genetics, any cis-acting element can be identified, including currently undescribed categories. Examples include sequences which contribute to (i) translation rate, (ii) polymerase binding sites which may operate during transcription within the virus core or during negative strand synthesis during virus assembly, (iii) microRNA binding sites, (iv) regulation of packaging and genome assortment and (v) RNA stability. Combining the mapping of lethal mutations with secondary structure prediction has permitted the identification of a region of potential secondary structure encompassing sequences from both ends of the segment 10 transcript, which contains all the functional cis-acting sequences in segment 10, within the limit of the spacing of the interrogated bases. These bases, nt 1–43 and nt 715–822, in combination with the predicted secondary structure can be used to inform further functional studies.

Fig. 3. Primary and secondary scans of the 3’ end of segment 10 mapped onto the predicted secondary structure. Labelling as described for Fig. 2. *, Positions within the predicted stem which tolerated variation.
Fig. 4. Primary and secondary scans of the 5′ and 3′ end regions of segment 10 mapped onto the predicted secondary structure. Labelling as described for Fig. 2. The circumscribed region of predicted secondary structure contains all the invariant positions identified.

relating to cis-acting sequences regulating virus-directed processes relating to segment 10.

The extreme ends of the segment could not be varied, in agreement with these being the initiation sites of both transcription and genome replication, which require the viral RNA-dependent RNA polymerase (RdRp) to bind to termini of the segment. The binding of the template by the RdRp and initiation of RNA synthesis in either mode of synthesis may require undefined sequence-specific or RNA-structure-specific interactions which do not tolerate variation. The UTRs of bluetongue virus segments have been shown to be differentially regulated at the level of translation of the ORF, depending on the segment from which they are derived, contributing to the highly varied levels of expression observed among the virus proteins (Boyce et al., 2012; Huismans, 1979). Within this observation it should be expected that RNA motifs which regulate either the rate at which the protein is synthesized or the stability of its mRNA would be included in the regions identified.

The degenerate mutagenesis approach used in this study could be applied to the termini of diverse viruses for which a PCR step can be incorporated into the reverse genetics system. A limiting consideration is how much degeneracy can be introduced into the PCR amplicon population at one time without reducing the abundance of viable genomes beyond the point where sufficient rescue events occur to generate a diverse population for the sequence analysis. In the bluetongue virus reverse genetics system, \( \sim 4 \times 10^2 \) to \( 8 \times 10^2 \) unique rescue events are recovered using \( 2 \times 10^6 \) cells, and with every essential position that is varied as a binary choice the number of viable rescue events may be expected to be reduced by 50%. The reduction observed in the number of plaques in the transfected monolayers is in agreement with this calculation, with five essential positions being varied producing the expected \( \sim 2^5 \) (32-fold) drop in virus recovery whilst still generating sufficient unique rescue events for sequence analysis. In other viral reverse genetics systems, with higher recovery rates, such as positive-sense RNA viruses where a single molecule is infectious, the number of bases varied could be increased or alternatively the number of mapping experiments required to generate a desired density of mutagenesis reduced correspondingly. This type of degenerate mutagenesis may be used in diverse virus systems, initially to identify functionally essential sequences of any type and secondly to probe the variation tolerated in the motifs identified. With oligonucleotide synthesis chemistry being capable of reliably generating primers of up to \( \sim 200 \) nt length, this approach may be applied to mapping across this length of region to identify essential motifs which are functional at the level of the viral genome, viral mRNA or the encoded protein.
Fig. 5. The viability of mutants with the base pairing within the predicted stem–loop either grossly disrupted or restored through the introduction of alternative base pairing. (a–e) The mutations made in the predicted stem, with arrows indicating the mutated positions in the predicted secondary structure of each mutant. Bars indicate the bases composing the predicted loop in the WT transcript. Minimum free energy secondary structure predictions generated using Mfold. Numbering indicates the nucleotide position within segment 10. (f–k) Virus-induced cytopathic effect visualized by crystal violet staining of the mutants in (a–e).
METHODS

**Cell lines.** BSR cells (a BHK-21 subclone) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% (v/v) FBS at 36 °C in 5% CO₂.

**Virus.** The South African reference strain of bluetongue virus serotype 1 was generated by reverse genetics (Boyce et al., 2008), and segment 10 variants were recovered by reverse genetics.

**Degenerate oligonucleotide primers.**

Degen 1a 39

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

Degen 1b 39

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

Degen 2a 39

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

Degen 2b 39

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

Underlined bases indicate positions containing sequence degeneracy denoted with IUPAC ambiguity codes. The introduced degeneracy was 50% of each base, producing an equal frequency of the WT base and the mutant base. The T7 promoter is shown in bold type, and the bluetongue sequence is in italics. Oligonucleotides were synthesized by Integrated DNA Technologies.

Degenerate primers were used in combination with WT primers S10 wt 5' or S10 wt 3' (see below) to produce degenerate templates for run-off transcription.

**Site-directed mutagenesis primers targeting the segment 10 predicted stem.**

S10 740 to 745 stem disrupt R

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 745 to 746 stem restore R

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 739 to 746 stem disrupt R

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 739 to 746 stem restore R

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 wt 3'

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 wt 5'

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

Mutagenic or WT reverse primers were used in combination with WT S10 wt 5' primer to produce templates for run-off transcription.

**Primers used in the sequence-independent amplification of the extreme termini of segment 10.**

Hairpin anchor primer 5' Phos GACCTCTGAGGATTCTAAAC/iSp9/ TCCAGITTTAGAATCC

‘/iSp9/’ denotes the nine-atom triethylene glycol spacer, adjacent to the hairpin.

FLAC2 Primer

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 392R GAGCAACAAACCGCCTATCG

S10 649F CAATCATATAATGATGCGGTGAGG

S10 564F CTTCCTGATGATGGTCTGCGC

S10 wt 3'

<table>
<thead>
<tr>
<th>GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA-</th>
<th>WCCACTCTACAC</th>
<th>SATGCAG-ACCTC-</th>
<th>SGGGCGCCACTC-</th>
<th>WACCTACTGATC-</th>
<th>YTAGGTTAATGG-</th>
<th>AATTCGAAACC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTCTAGCGGGACTGATGAAAATTCC</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
</tbody>
</table>

S10 564F CTTCCTGATGATGGTCTGCGC

S10 wt 3' GTAAGTGTGTAAGCCGCTACATCCCTCCCCGTATAGACAGCA- WCCACTCTACAC | SATGCAG-ACCTC- | SGGGCGCCACTC- | WACCTACTGATC- | YTAGGTTAATGG- | AATTCGAAACC- |

| RTCTAGCGGGACTGATGAAAATTCC |------------------|------------------|------------------|------------------|------------------|------------------|

**Synthesis of viral single stranded RNAs (ssRNAs).** Synthetic ssRNAs were prepared by run-off in vitro transcription, using T7 RNA polymerase either from plasmid clones which had been linearized
with BsuBI, 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 templates using the mMESSAGE mMACHINE T7 Ultra kit (Ambion), as previously described (Boyce et al., 2008).

**Transfection of BSR cells to generate mutant bluetongue virus by reverse genetics.** The variants of segment 10 were rescued in the bluetongue virus 1 South African reference strain by transfection of the BSR cell line, as previously described (Boyce et al., 2008). Variants of the segment 10 transcript were generated from the variant T7 PCR amplicons produced using a mutant reverse/forward primer with the WT forward/reverse primer. Transfected monolayers were overlaid with Modified Eagle Medium (MEM), 1.5 % agarose type VII-A (Sigma-Aldrich), supplemented with 5 % (v/v) FBS at 36 °C in 5 % CO₂ to permit the number of plaques generated with each degenerate segment 10 transcript population to be counted. A minimum diversity threshold of 20 rescued plaques was set for continuing with amplification of the recovered population.

**Amplification of progeny viruses.** The populations of progeny viruses recovered were titrated on BSR monolayers and overlaid with MEM, 1.5 % agarose type VII-A (Sigma-Aldrich), supplemented with 5 % (v/v) FBS at 36 °C in 5 % CO₂. Individual plaques were picked into 500 µl DMEM supplemented with 8 % (v/v) FBS and 200 µl of each picked plaque was propagated on 80 % confluent BSR monolayers in 12-well plates at 36 °C in 5 % CO₂ until the cytopathic effect (CPE) included > 70 % of the cells (3–4 days).

**RT-PCR amplification of the target region of progeny viruses.** Total RNA was extracted from amplified progeny viruses, using TRIReagent (Sigma) according to the manufacturer’s instructions. dsRNA was purified by selective precipitation of the single-stranded RNA component in 2 M LiCl at 4 °C for 16 h. The soluble dsRNA fraction was collected following centrifugation at 15 000 g at 4 °C for 10 min and precipitated using 2-propanol with 0.3 M sodium acetate (pH 5.2). The dsRNA was pelleted by centrifugation at 15 000 g at 4 °C for 10 min, washed in 70 % (v/v) ethanol and dissolved in nuclease-free water. To permit amplification of the extreme termini of the segment, a hairpin anchor primer was ligated to the 3’ ends of the dsRNA segments using T4 RNA ligase, as described by Maan et al., 2008. Total RNA was extracted from amplified progeny viruses, using TRIReagent (Sigma) according to the manufacturer’s instructions. dsRNA was purified by selective precipitation of the single-stranded RNA component in 2 M LiCl at 4 °C for 16 h. The soluble dsRNA fraction was collected following centrifugation at 15 000 g at 4 °C for 10 min and precipitated using 2-propanol with 0.3 M sodium acetate (pH 5.2). The dsRNA was pelleted by centrifugation at 15 000 g at 4 °C for 10 min, washed in 70 % (v/v) ethanol and dissolved in nuclease-free water. To permit amplification of the extreme termini of the segment, a hairpin anchor primer was ligated to the 3’ ends of the dsRNA segments using T4 RNA ligase, as described by Maan et al., 2008. cDNA copies were made using the FLAC2 primer which anneals to the hairpin anchor primer, using RevertAid Premium reverse transcriptase (Thermo Scientific). The target region was amplified by PCR with the FLAC2 primer and either the S10 564F or S10 392R internal segment 10 primers, depending on which end of the segment was targeted, using Taq DNA polymerase. Each PCR amplicon was purified using a silica column (Promega Wizard SV Gel and PCR clean-up column) to separate it from residual primers, prior to sequencing.

**Sequence analysis of progeny viruses.** PCR amplicons were sequenced using an ABI 3730 DNA sequencer (Applied Biosystems). The identity of the varied positions in each progeny virus was determined from the electropherograms.

**Prediction of intramolecular secondary structures.** The prediction of intramolecular base pairing by minimum free energy was performed with the full-length positive-sense strand of the complete genome segment, using Mfold (Zuker, 2003). Predicted secondary structures were visualized using VARNA (Darty et al., 2009).

**Visualization of viral CPE caused by segment 10 stem mutants.** At 4 days following the transfection used for reverse genetics, the monolayers were washed in Hank’s Balanced Salt Solution (HBSS, ThermoFisher), and fixed with PBS, 10 % formaldehyde 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.

**REFERENCES**


