Nucleotide sequence shows that Bean leafroll virus has a Luteovirus-like genome organization

Leslie L. Domier,1, 2 Nancy K. McCoppin,1 Richard C. Larsen3 and Cleora J. D’Arcy2

1, 2United States Department of Agriculture, Agricultural Research Service1, and Department of Crop Sciences2, 1102 S. Goodwin Ave, University of Illinois at Urbana/Champaign, Urbana, IL 61801, USA
3 United States Department of Agriculture, Agricultural Research Service, 24106 N Bunn Rd, Prosser, WA 99350-8694, USA

The complete nucleotide sequence of the Bean leafroll virus (BLRV) genomic RNA and the termini of its smallest subgenomic RNAs were determined to better understand its mechanisms of gene expression and replication and its phylogenetic position within the Luteoviridae. The number and placement of open reading frames (ORFs) within the BLRV genome was Luteovirus-like. The nucleotide and predicted amino acid sequences of BLRV were most similar to those of Soybean dwarf virus (SbDV). Phylogenetic analyses employing the neighbour-joining method and sister-scanning analysis indicated that the BLRV nonstructural proteins were closely related to those of Barley yellow dwarf virus-PAV (BYDV-PAV), a Luteovirus. The region surrounding the frameshift at the junction between ORFs 1 and 2 also contained sequences very similar to those of BYDV-PAV and a Dianthovirus, Red clover necrotic mosaic virus. Similar analyses showed that the structural proteins were most similar to those of the Polerovirus genus. The 3’-noncoding regions downstream of ORF5 contained sequences similar to translational control elements identified in the BYDV-PAV genome. These data suggest that BLRV, like SbDV, is derived either through selection from a common ancestor with BYDV-PAV or that BLRV is the product of two recombination events between luteovirus-like and polerovirus-like ancestors where the 5’ 2900 nt and 3’ 700 nt of the BLRV genome are from a Luteovirus and the intervening sequences are derived from a Polerovirus.

Introduction

Bean leafroll virus (BLRV) is a member of the Luteoviridae family and infects several species of annual and perennial legumes, including many cultivated bean and pea species (Ashby, 1984). BLRV is distributed worldwide and has been found to be endemic in white clover in the southeastern United States (Damsteegt et al., 1995). Like other members of the Luteoviridae, BLRV is transmitted obligately by aphids in a persistent manner (Ashby, 1984).

The Luteoviridae family has been divided into three genera (Enamovirus, Luteovirus and Polerovirus) depending on genome organization, sequence similarity, and methods of gene expression (D’Arcy et al., 2000). The single-stranded, positive-sense genomes of the Luteoviridae contain five to six open reading frames (ORFs) designated ORF 0 through ORF 6. ORF 0 is unique to the Enamovirus and Polerovirus genera and encodes a protein of unknown function. ORFs 1 and 2 encode the replication-related proteins, which in luteoviruses are most similar to those of the Tombusviridae, while the replicases encoded by poleroviruses and enamoviruses are related to those of the Sobemovirus genus (D’Arcy et al., 2000). In all three genera, ORF 2 is expressed via a translational frameshift from ORF 1. ORF 1 overlaps ORF 2 by less than 20 nt in luteoviruses, but by more than 400 nt in enamo- and poleroviruses. The intergenic region between ORFs 2 and 3 is about 100 nt in luteoviruses and about 200 nt in polero- and enamoviruses. ORFs 3 and 5 encode the coat and readthrough proteins of the viruses. Recently, parallels have been found between the coat proteins (CPs) of poleroviruses and members of the Sobemovirus genus (Terradot et al., 2001), which suggests an even closer affiliation of the two genera. ORF 4, which is lacking in the enamoviruses, putatively encodes a movement protein.

Comparisons of the nucleotide and predicted amino acid sequences of members of the Luteoviridae with other members of the family and different virus families suggest that RNA
recombination has played an important role in the generation of new species within the Luteoviridae (Miller et al., 1995). Recombination seems to have occurred most commonly near the site of initiation of synthesis of the subgenomic RNA (sgRNA) that encodes the capsid and movement proteins. As a consequence, the Luteoviridae contains species with similar structural proteins, but disparate RNA-dependent RNA polymerases (RdRps). For example, Cucurbit aphidborne yellow virus (CABYV) is thought to be derived from recombination between poler- and enamo-like viruses (Gibbs et al., 2000; Gibbs & Cooper, 1995). Similarly, Sugarcane yellow leaf virus (ScYLV) is thought to have arisen by recombination between polero- and enamo-like viruses (Smith et al., 2000). Finally, it has been suggested that SbDV is derived from recombination events between luteo- and poler-like viruses (Rathjen et al., 1994; Terauchi et al., 2001). Since one of the primary distinguishing features of the three Luteoviridae genera is the type of RdRp encoded by the virus, CP sequences alone are not sufficient to assign a virus species to a genus.

The relationship of BLRV to other members of the Luteoviridae has been examined based on biological, immunological, nucleic acid hybridization and nucleotide sequence data. Serological (D’Arcy et al., 1988; Smith et al., 1996; van den Heuvel et al., 1990) and nucleic acid hybridization data (Martin & D’Arcy, 1990) suggest that BLRV is related most closely to Soybean dwarf virus (SbDV). Sequences of the BLRV coat protein gene have been reported (Cavileer & Berger, 1994; Prill et al., 1990) and predicted that BLRV was most closely related to poleroviruses. As mentioned above, SbDV has a genome organization similar to members of the genus Luteovirus, but may have arisen through recombination between a Luteovirus and a Polerovirus (Rathjen et al., 1994; Smith et al., 2000; Terauchi et al., 2001). In this study we determined the complete genomic sequence of a Michi- can isolate of BLRV and compared the coding and noncoding regions of the genomic sequence to determine to which of the three Luteoviridae genera BLRV was most similar.

Methods

- **Viruses isolate.** A Washington isolate of BLRV was maintained in broad bean (Vicia faba L.) by transmission with Aphis gossypii (Harris). A Michigan BLRV isolate (PV-2651) was obtained from the ATCC in lyophilized broad bean tissue. The genomic nucleotide sequence was derived from the Michigan (ATCC) BLRV isolate. The terminus of the smallest subgenomic RNA was determined by using the Washington BLRV isolate.

- **Sequence analysis.** Total RNA was extracted from leaf and lyophilized broad bean tissue using Trizol reagent according to the manufacturer’s protocols (Invitrogen). To amplify regions of the BLRV genome downstream and upstream of the previously published CP gene (nt 3093–3683; Cavileer & Berger, 1994; Prill et al., 1990), degenerate primers were designed from conserved sequences in the SbDV and BYDV-PAV genomes. A DNA fragment representing nt 1276–3151 was amplified using 5′ CGKTTTTAGAGGGCCTCG 3′ (nt 1276–1295) and a primer complementary to nt 3131–3151. A fragment representing nt 317–1355 was amplified with 5′ CAYGAYGCYTTGTGSCACATG-TGCT 3′ (nt 317–341) and a primer complementary to nt 1340–1355. A fragment containing nt 3574–4401 was amplified using 5′ TTCAART-GCTCATCAAG 3′ (nt 4324–4349) and a primer complementary to nt 456–472. For the dGTP-tailed cDNA, fragments containing the 5′ terminus of the BLRV genome were amplified using an oligo(dT) primer [5′ CGACCTGTACCTGACAT(C)3′] and a primer complementary to nt 456–472. For the dGTP-tailed cDNA, fragments containing the 5′ terminus of the BLRV genome were amplified in nested PCR reactions using an oligo(dC) primer [5′ CGACCTGTACCTGACAT(C)3′] and a primer complementary to nt 322–345. To determine the sequence of the 5′ terminal region of the BLRV genome (nt 4324–5905), total RNA extracted from BLRV-infected broad bean tissue was polyadenylated using poly(A) polymerase according to the manufacturer’s recommendations (Promega). Because the genomic sequences of SbDV and BYDV-PAV contain at least three 3′-terminal C residues, an anchored oligo(dT) primer [5′ GGCGTTGGGTAGCTAC(T)3′] was used to synthesize cDNA and amplify the 5′-terminal fragment in combination with a primer corresponding to positions nt 4324–4349.

Blends of the BLRV sgRNAs were predicted based on nucleotide sequence similarity to the 5′ terminus of genomic RNA. To determine the 5′ terminus of the largest sgRNA, cDNA was primed with an oligonucleotide complementary to nt 3125–3160, tailed with dGTP and amplified in nested PCR reactions with the oligo(dC) primer and oligonucleotides complementary to nt 3027–3061 and 2962–2995. To determine the 5′ terminus of the intermediate sgRNA, cDNA was primed with an oligonucleotide complementary to nt 5395–5422, tailed with dGTP and amplified in nested PCR reactions with the oligo(dC) primer and oligonucleotides complementary to nt 5281–5304 and 5197–5233. To determine the 5′ terminus of the smallest sgRNA, cDNA was primed with an oligonucleotide complementary to nt 5950–5965, tailed with dGTP and amplified in nested PCR reactions with the oligo(dC) primer and oligonucleotides complementary to nt 5890–5922 and 5835–5858.

- **Phylogenetic analysis.** Phylogenetic relationships of BLRV were inferred by comparing the predicted amino acid sequences of ORF 2 and ORF 3 to those of members of the Luteoviridae. Sequences were aligned using ClustalX (Thompson et al., 1997). Trees were constructed using the neighbor-joining method in PAUP version 4.0b6 (Sinauer Associates, Sunderland, MA, USA). Sister scanning analysis was performed using siccan version 2.0 (Gibbs et al., 2000) and the BYDV-PAV (AF235167) and PLRV (NC 001747) genomic sequences. A window of 100 nt was shifted 50 nt at each step. To allow scoring of the 5′ and 3′ noncoding regions, gapped positions were evaluated.

Results and Discussion

The phylogenetic position of BLRV was evaluated by comparing the nucleotide and predicted amino acid sequences, organization of ORFs and conserved regulatory elements with those of other members of the Luteoviridae and related viruses.
The nucleotide sequence of the BLRV genome (GenBank accession no. AF441393) was 5964 nt in length and revealed a genomic organization containing five large ORFs (Fig. 1) that was similar to that of members of the Luteovirus genus. There was no evidence of an ORF 0, which is found only in poleroviruses and enamoviruses. Like those of the SbDV isolates analysed (Rathjen et al., 1994; Terauchi et al., 2001), the BLRV sequenced lacked an ORF 6, which is found in most BYDVs. The largest ORF downstream of ORF 5 could encode a protein with a predicted mass of just 2.7 kDa (nt 5628–5696), much smaller than the 6.7 kDa proteins predicted from BYDV-PAVs. It should be noted that the sizes and sequences of ORFs 6s among BYDVs can be highly variable and that ORF 6 expression is not required for virus infectivity (Chalhoub et al., 1994; Miller et al., 1995). ORFs 1 and 2 overlapped by just 15 nt compared to over 400 nt for members of the Polerovirus and Enamovirus genera. The conceptual translation of ORFs 1 and 2 was predicted to encode a tombusvirus-like RdRp. Like poleroviruses, ORFs 2 and 3 were separated by 211 nt. This is similar to SbDV where the intergenic region is 208 nt (Rathjen et al., 1994; Terauchi et al., 2001). Luteoviruses typically have intergenic regions of about 100 nt, while poleroviruses possess intergenic regions of more than 200 nt. Hence, like SbDV, the BLRV genome was composed of luteovirus- and polerovirus-like elements.

**Noncoding regions**

The 5′ terminus of the genomic RNA was determined twice using dGTP and dATP in the tailing reactions. In both cases, the predicted terminus of the genomic RNA was the same. Even though reverse transcriptase has been reported to sometimes incorporate extra nucleotides at the 3′ end of the cDNA product (Chen & Patton, 2001), no evidence of non-template nucleotide addition was observed. BLRV genomic RNA had a predicted 5′ noncoding region (NCR) of 178 nt, longer than that of BYDV-PAV, and a 3′ NCR of 703 nt, shorter than that of BYDV-PAV. Nucleotides 2 through 21 of the BLRV genome were identical to the corresponding region of the SbDV genome. A region of similarly high sequence identity between BLRV and SbDV was located at their 3′ termini, corresponding to nt 5926–5965 of BLRV. These conserved regions may be important in replication and/or gene expression.

Even though the 3′ NCR of BYDV-PAV does not encode protein products, it is very important for the expression of viral proteins. The BYDV-PAV 3′ NCR contains a translational enhancer (3′TE) that mediates the cap-independent translation initiation of its genomic and subgenomic RNAs. The 3′TE binds sequences near the 5′ termini of translated BYDV-PAV RNAs and eukaryotic initiation factors eIF4F to 40S ribosomal subunits for translation initiation (Wang et al., 1997). The BYDV-PAV 3′ NCR also contains sequences that are required for the frameshifting between ORFs 1 and 2 (Paul et al., 2001). The 3′ NCR of BLRV (and that of SbDV) contained sequence elements similar to both of the BYDV-PAV control sequences. A region (nt 5293–5387) just downstream of ORF 5 in the BLRV nucleotide sequence showed a high degree of sequence identity and conserved secondary structure to the BYDV-PAV 3′TE (Fig. 2A). The longest consecutive set of conserved nucleotides (5′ AUCCUGGGAAAACAGG 3′) was shared by BLRV, SbDV and the BYDVs and was predicted to form the first stem-and-loop of the enhancer. The second and probably more interesting conserved structure was the third stem and loop which, like BYDV-PAV, contained a sequence complementary to regions in the 5′ NCR that were also capable of forming two small stem-and-loops structures (nt 2–20 and 30–51). These structures were similar to those recently proposed for SbDV (Guo et al., 2001). In BYDV-PAV, the interaction of these two regions is thought to mediate translation enhancement. These observations suggest that, like BYDV-PAV, the BLRV genes are expressed through cap-independent processes that are facilitated by the long-distance interaction of the 5′ and 3′NCRs.

Just downstream of the region similar to the BYDV-PAV 3′TE, the BLRV nucleotide sequence contained a region similar to the 3′frameshift signal (3′FSS) of BYDV-PAV (Fig. 2B). Paul et al. (2001) determined that the BYDV-PAV 3′FSS contained a domain that was required for frameshifting and a region that enhanced frameshifting. Surprisingly, the region that enhanced frameshifting was more highly conserved than the region required for frameshifting. This may be due to differences in

![BLRV nucleotide sequence](image)

Fig. 1. Diagram of the genome organization of BLV. ORFs are indicated by open boxes that are staggered vertically to indicate the relative reading frame occupied by each ORF. The positions of sequences similar to the BYDV-PAV 3′-translation enhancer (3′TE), 3′-frameshift signal (3′FSS) and nucleotides complementary to the 3′-proximal unpaired loop of the putative BLRV 3′TE are indicated.
sequences with which the element must interact in the different viruses.

**Intergenic region**

Even though the predicted amino acid sequences of proteins derived from ORFs 1 and 2 were very similar to those of BYDV-PAV, the length of the intergenic region between ORFs 2 and 3 of BLRV (211 nt) and the predicted amino acid sequences of the protein products of ORFs 3 and 5 were more similar to those of the poleroviruses. Other studies have mapped recombination events to this area (Chalhoub & Lapierre, 1995; D’Arcy et al., 2000; Gibbs & Cooper, 1995; Smith et al., 2000). It has been proposed that the mechanism underlying the initiation of transcription of the sgRNA from this position might favour template switching for RNA recombination (Miller et al., 1995). This size difference would be consistent with those proposals. To generate the BLRV genome organization, a second recombination event near the end of ORF 5 would have had to occur to place a luteovirus-like 3′ NCR downstream of ORF 5. As proposed for the recombination event between ORFs 2 and 3, the recombination could have been facilitated by the initiation of RNA synthesis at the sgRNA promoter located in this region in Luteovirus genomes.

**ORFs 1 and 2**

ORF 1 (nt 179–1285) and ORF 2 (nt 1255–2883) were predicted to encode proteins of 42 kDa and 62 kDa, respectively. Their predicted amino acid sequences were most similar to those of SbDV, followed by BYDV-PAV. BLAST comparisons with BLRV ORF 1 yielded E values of $e^{-132}$ (62% identity) with SbDV and $4e^{-35}$ (27% identity) with BYDV-PAV. ORF 2 was more conserved and yields E values of 0 with both SbDV (82% identity) and BYDV-PAV (62%). There were no significant alignments with members of the Polerovirus or Enamovirus genera. As with other members of the Luteoviridae, ORF 2 is predicted to be translated as a translational fusion with the product of ORF 1 through a frameshift. In many systems, frameshifting is promoted by a ‘slippery’ sequence that allows frameshifting ribosomes to retain the tRNAs in the peptidyl and aminoacyl sites during the −1 transition (Jacks et al., 1988). A sequence which fits the canonical slippery sequence was found at the junction of BLRV ORFs 1 and 2 (GGUUUUC); this is identical to that of SbDV.
and similar to that of BYDV-PAV (GGGUUUU) and Red clover necrotic mosaic virus (RCNMV; GGAUUUU), a Dianthovirus in the Tombusviridae (Kim & Lommel, 1998). Like BYDV-PAV, the shifty sequence was flanked by sequences that could form two bulged stem-and-loop structures. Comparisons of the sequences and predicted secondary structures showed that the upstream structure was highly variable. The upstream structure was not required for frameshifting of RCNMV (Kim & Lommel, 1998). In contrast, the downstream structure was highly conserved among BLRV, BYDVs, SbDV and RCNMV (Fig. 2C). Thirty-one nucleotides downstream of the GGDUUUU slippery sequence and UAG termination codon was the sequence CCCVUWWYCUAUUYUCSG, which was conserved in all the viruses and formed a loop at the top of the second stem. All of the predicted structures also contained bulges with the sequence UUGA unpaired on the upstream side of the stem. It has been suggested that two features, a shifty sequence and a complex secondary structure downstream of the termination codon, are needed to promote efficient frame shifting. All of the viruses here have both elements. However, the primary sequence conservation in the unpaired bulge and loop of the downstream structure suggests that these sequences may interact with a second factor, possibly conserved host proteins or RNAs.

**ORFs 3, 4, and 5**

BLRV ORFs 3 (nt 3093–3683), 4 (nt 3238–3664) and 5 (nt 3684–5262) were predicted to produce proteins of 22, 16 and 59 kDa, respectively. In contrast to ORFs 1 and 2, the predicted amino acid sequences of BLRV ORFs 3, 4 and 5, after SbDV, were most similar to members of the Polerovirus genus. BLAST comparisons using the BLRV CP amino acid sequence, encoded by ORF 3, produced their highest E values with SbDV (e<sup>−13</sup>, 72% identity) followed by Groundnut rosette assist virus (GRAV; e<sup>−46</sup>, 60% identity) and Potato leafroll virus (PLRV; e<sup>−40</sup>, 52% identity). ORF 4 was the least conserved of the five BLRV ORFs, and yielded E values with SbDV of 2e<sup>−13</sup> (37% identity), with Beet western yellow virus of e<sup>−10</sup> (34% identity) and with PLRV of 2e<sup>−48</sup> (32% identity). Finally, comparisons with ORF 5 yielded E values with SbDV of 2e<sup>−49</sup> (57% identity), with Beet mild yellow virus of 2e<sup>−83</sup> (62% identity) and with BYDV-RPV of 4e<sup>−82</sup> (54% identity). For comparison, E values for alignments with BYDV-PAV for ORFs 3 and 5 produced E values of 4e<sup>−43</sup> and 8e<sup>−48</sup>, respectively. No significant homology was reported between ORF 4s of BLRV and BYDV-PAV. These results showed that the E values in the CP regions of BLRV, PLRV and BYDV-PAV were similar. However the E values derived from the ORF 5 comparisons showed a much closer relationship between BLRV and the poleroviruses.

The above results are consistent with observations from immunosorbet electron microscopy that showed strong interactions between polyclonal antisera for BLRV and GRAV (Casper et al., 1983; Reddy et al., 1985). Recently the three-dimensional structure of the PLRV CP was predicted based on homology modelling by aligning the PLRV sequence with that of Rice yellow mottle virus, a member of the Sobemovirus genus (Terradot et al., 2001). The analysis predicted that five acidic amino acids (E<sub>109</sub>, E<sub>117</sub>, D<sub>172</sub>, E<sub>176</sub> and D<sub>177</sub>) are present on the surface of PLRV particles at the trimer axis. These five residues are conserved among the poleroviruses (Fig. 3). However, all of the monocot-infecting members of the Luteovirus genus (BYDVs PAV, PAV129, MAV and SGV) and ScYLV have hydrophobic or polar substitutions at E<sub>176</sub> (Fig. 3). The BLRV sequence, like that of SbDV, contained the conserved E<sub>109</sub>, D<sub>172</sub>, E<sub>176</sub> and D<sub>177</sub>, but contained a hydrophobic substitution (A) at the E<sub>176</sub> position. The amino acid downstream of E<sub>109</sub> were conserved in the polerovirus and luteovirus lineages of

---

**BLRV sgRNAs**

<table>
<thead>
<tr>
<th>sgRNA</th>
<th>BLRV nucleotide sequence</th>
</tr>
</thead>
</table>
| **1** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **2** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **3** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **4** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **5** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence

---

**BLRV sgRNAs**

<table>
<thead>
<tr>
<th>sgRNA</th>
<th>BLRV nucleotide sequence</th>
</tr>
</thead>
</table>
| **1** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **2** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **3** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **4** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **5** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence

---

**BLRV sgRNAs**

<table>
<thead>
<tr>
<th>sgRNA</th>
<th>BLRV nucleotide sequence</th>
</tr>
</thead>
</table>
| **1** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **2** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **3** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **4** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **5** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence

---

**BLRV sgRNAs**

<table>
<thead>
<tr>
<th>sgRNA</th>
<th>BLRV nucleotide sequence</th>
</tr>
</thead>
</table>
| **1** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **2** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **3** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **4** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **5** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence

---

**BLRV sgRNAs**

<table>
<thead>
<tr>
<th>sgRNA</th>
<th>BLRV nucleotide sequence</th>
</tr>
</thead>
</table>
| **1** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **2** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **3** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **4** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
| **5** | 1GUGAAAG---UUGACUCCUUCACAAAGGCUG---UUGACUGCUUACAG | BLRV nucleotide sequence
the Luteoviridae. However, these residues were much less conserved in the CP sequences of BLRV, SbDV, Sweet potato leaf speckling virus (SPLSV), SCYLV and Pea enation mosaic virus 1 (PEMV-1). The viruses in this group did not fit neatly into either the Luteovirus or Polerovirus genera. PEMV-1 is the sole member of the Enamovirus genus. As mentioned above, BLRV, SbDV and SCYLV seem to have arisen through interspecific recombination. (Smith et al., 2000; Rathjen et al., 1994; Terauchi et al., 2001). Because only the CP sequence has been reported for SCYLV (Fuentes et al., 1996), its genomic organization is not known. While recombination may have played a role in the generation of these genomes, the lack of conservation in the amino acids thought to be on the surface of the virions may reflect unique selection pressures placed on the viruses by their hosts and/or aphid vectors.

Termini of subgenomic RNAs

Another feature that distinguishes luteo- and poleroviruses is the number of sgRNAs produced. While some poleroviruses produce two sgRNAs (Ashoub et al., 1998), BYDV-PAV synthesizes three sgRNAs (Miller et al., 1995). The termini and promoters of the BYDV-PAV sgRNA have been mapped (Koev & Miller, 2000; Koev et al., 1999). Because 5’ termini of sgRNAs coronaviruses that produce multiple sgRNAs often contain sequences similar to the genomic RNA (Lai & Cavanagh, 1997), the BLRV genome was analysed for sequences similar to the 5’ terminus of its genomic RNA. Sequences identical to nt 2 through 7 (GUAAG) of the genomic RNA were found at positions 2792, 5036 and 5646. The first was located 302 nt upstream of the AUG for ORF 3, which encodes CP. The second was located at position 5260, just inside the 3’ end of ORF 5. The third was located 319 nt from the 3’ terminus of BLRV genomic RNA. The last two positions were very similar to the positions of the two small subgenomic RNAs produced by BYDV-PAV (Dinesh-Kumar et al., 1992; Kelly et al., 1994; Koev & Miller, 2000). Northern blot analysis of total RNA extracted from BLRV-infected broad bean leaves detected only the presence of the small sgRNA of about 300 nt (data not shown). Analysis of the sgRNA by 5’ RACE using oligonucleotide primers downstream of the predicted termini produced a band of the expected size for the smallest sgRNA. Nucleotide sequence analysis of the amplified fragment showed that the sgRNAs terminated at position 5646, as predicted from the sequence comparison. Specific bands were not obtained for the two larger sgRNAs, possibly due to the low abundance or labile nature of the RNAs. Smith et al. (1991) showed that dsRNA preparations from SbDV-infected plant tissue contained only genomic RNA and large sgRNAs. They found no evidence of smaller sgRNAs. Unlike BLRV, the SbDV genome contains only two internal copies of the GUAAG sequence, one that would produce the large sgRNA and one that would produce a small (about 300 nt) sgRNA (Fig. 4). It is possible that SbDV produces only two sgRNAs, one of which was detected in the analysis of dsRNA and one that was not retained on the gels. Even though only one of the sgRNAs was detected, its size was similar to the smallest sgRNA produced by BYDV-PAV. Additional studies will be required to determine whether BLRV and SbDV produce intermediate-sized sgRNAs.

Relationships to other viruses

Phylogenetic analysis suggests that SbDV and BLRV were derived from the same recombination or speciation events and that they evolved independently from that point (Fig. 5).
Hence the recombination events would have occurred before the speciation of SbDV and BLRV. Phylograms produced from the ORF 2 predicted amino acid sequence positioned BLRV and SbDV in the Luteovirus genus (Fig. 5A). A phylogram constructed from CP predicted amino acid sequences showed that BLRV and SbDV formed a separate clade that branched near the base of the tree with only a slightly stronger association of BLRV and SbDV with the poleroviruses than the luteoviruses (Fig. 5B). This conclusion is supported by Sister-scanning analysis using BLRV, BYDV-PAV and PLRV nucleotide sequences, which showed that the BLRV and BYDV-PAV sequences were least similar in ORFs 3 and 5 (Fig. 5C). While the PLRV sequence was most similar to that of BLRV in this region, the BYDV-PAV sequence produced some Z scores nearly as high. Using the current criteria for assigning a virus species to a genus within the Luteoviridae BLRV and SbDV would be placed with the Luteovirus genus. However, while BLRV and SbDV possess luteovirus-like genome organizations and presumably employ luteovirus-like replication and expression strategies, they encode polerovirus-like structural protein. Hence, these criteria do not capture the genetic diversity present within viruses that may have arisen through recombination.

References


Received 15 January 2002; Accepted 26 February 2002