Nucleotide sequence of RNA2 of *Lettuce big-vein virus* and evidence for a possible transcription termination/initiation strategy similar to that of rhabdoviruses

Takahide Sasaya,1 Shinnosuke Kusaba,1 Koichi Ishikawa1 and Hiroki Koganezawa2

1National Agricultural Research Center for Western Region, Shikoku Campus, Zentsuji, Kagawa 765-8508, Japan
2National Agricultural Research Center for Western Region, Fukuyama, Hiroshima 721-8514, Japan

*Lettuce big-vein virus* (LBVV) is the type species of the genus *Varicosavirus* and is a two-segmented negative-sense single-stranded RNA virus. The larger LBVV genome segment (RNA1) consists of 6797 nt and encodes an L polymerase that resembles that of rhabdoviruses. Here, the nucleotide sequence of the second LBVV genome segment (RNA2) is reported. LBVV RNA2 consisted of 6081 nt and contained antisense information for five major ORFs: ORF1 (nt 210–1403 on the viral RNA), ORF2 (nt 1493–2494), ORF3 (nt 2617–3489), ORF4 (nt 3843–4337) and ORF5 (nt 4530–5636), which had coding capacities of 44, 36, 32, 19 and 41 kDa, respectively. The gene at the 3' end of the viral RNA encoded a coat protein, while the other four genes encoded proteins of unknown functions. The 3'-terminal 11 nt of LBVV RNA2 were identical to those of LBVV RNA1, and the 5'-terminal regions of LBVV RNA1 and RNA2 contained a long common nucleotide stretch of about 100 nt. Northern blot analysis using probes specific to the individual ORFs revealed that LBVV transcribes monocistronic RNAs. Analysis of the terminal sequences, and primer extension and RNase H digestion analysis of LBVV mRNAs, suggested that LBVV utilizes a transcription termination/initiation strategy comparable with that of rhabdoviruses.

**INTRODUCTION**

*Lettuce big-vein* is a fungus-borne disease that is a serious problem in many lettuce-growing areas. *Lettuce big-vein virus* (LBVV), which is almost always isolated from lettuce plants (*Lactuca sativa* L.) showing big-vein symptoms, has been considered the causal agent of the disease for nearly two decades (Kuwata *et al*., 1983; Vetten *et al*., 1987; Huijberts *et al*., 1990). However, a second virus, Mirafiori lettuce virus in the genus *Ophiovirus*, was recently found in lettuce showing big-vein symptoms, so the exact role of LBVV in big-vein disease remains unclear (Roggero *et al*., 2000; Lot *et al*., 2002; Natsuaki *et al*., 2002). LBVV is the type species of the genus *Varicosavirus*, which also includes Tobacco stunt virus, Camellia yellow mottle virus and Freesia leaf necrosis virus (Mayo, 2000). LBVV virions are rod-shaped with modal lengths of 320–360 nm and diameters of about 18 nm. LBVV virions contain a single coat protein (CP) with an *M*<sub>r</sub> of 48 000 and do not have an envelope structure (Kuwata *et al*., 1983). LBVV was reported to be a two-segmented double-stranded RNA (dsRNA) virus, which contained two dsRNAs of approximately 7-0 and 6-5 kbp (Mayo, 2000). However, we recently demonstrated that LBVV is not a dsRNA virus but a single-stranded negative-sense RNA virus with a bipartite genome. Viral (negative-sense) and virus-complementary (positive-sense) RNAs are separately encapsidated in the virions and the viral RNA is predominant (Sasaya *et al*., 2001, 2002).

LBVV was found in lettuce plants showing big-vein symptoms in the early 1980s (Kuwata *et al*., 1983; Vetten *et al*., 1987; Huijberts *et al*., 1990), but its molecular characterization has been slow because LBVV particles are refractory to standard methods for virus purification due to their instability and tendency to aggregate. After numerous efforts, we succeeded in obtaining purified LBVV virions from lettuce plants and determining the nucleotide sequence of the CP gene on RNA2 and the L polymerase gene (designated by analogy with the L polymerase of rhabdoviruses, which is a catalytic subunit of the RNA polymerase).
complex) on RNA1 (Sasaya et al., 2001, 2002). Analysis of amino acid sequences of the LBVV CP and L polymerase suggested a close relationship between LBVV and rhabdoviruses in the order Mononegavirales, even though LBVV has a two-segment negative-sense RNA genome and its particles are not enveloped. In particular, the LBVV L polymerase is most closely related to that of Northern cereal mosaic virus (NCMV) in the genus Cytorhabdovirus (Sasaya et al., 2002). In this report, we present the nucleotide sequence of LBVV RNA2 and the predicted non-coding and coding regions. We also determined the 5′ and 3′ ends of LBVV mRNAs to get an insight into the possible transcription mechanism of LBVV. Our data provide further evidence for a relationship between LBVV and rhabdoviruses, as well as illustrating some unique properties of LBVV.

**METHODS**

**Virus and RNA purification.** The source of LBVV, procedure for virus purification and isolation of viral RNAs have been described previously (Sasaya et al., 2001, 2002). Poly(A)+ RNA was isolated from healthy and LBVV-infected lettuce plants 4 weeks after transplanting to stream-sterilized and LBVV-infested soils, respectively, using a Dynabeads mRNA Direct kit, following the instructions provided by the supplier (Dynal).

**Cloning strategy for LBVV RNA2.** A cDNA library of LBVV RNA2 was constructed and its sequence was determined by a genome-walking method, as described by Fazeli & Rezaian (2000). The cloning strategy for LBVV RNA2 and the synthetic oligonucleotide primers used for the cloning are shown in Fig. 1 and Table 1. Briefly, first-strand cDNA synthesis was done using 20 ng purified LBVV genomic RNAs primed with the LBVV-specific first-strand cDNA synthesis primer 1st GW, whose sequence corresponded to nt 1135–1151 of the previously described LBVV CP gene (Sasaya et al., 2001). The second-strand cDNA synthesis primer P1-N6, which consisted of a random hexamer linked to a known oligonucleotide at its 5′ end, was added to the first-strand reaction mixture and primer extension was carried out at 37°C for 30 min in a 50 μl reaction volume containing 50 mM Tris/HCl, pH 7.2, 10 mM MgSO4, 0.1 mM DTT, 0.3 mM each of the four dNTPs and 5 U Klenow fragment (Toyobo). After excess oligonucleotide had been removed by centrifugal ultrafiltration (Centricon 100 microconcentrators; Amicon), double-strand cDNA was amplified by PCR with the known oligonucleotides P1 and the first LBVV-specific genome-walking primer, GW-1p. The PCR product was cloned and sequenced as described previously (Sasaya et al., 2002). The next region of LBVV RNA2 was amplified by PCR using the P1 primer and a new LBVV-specific primer, which was designed on the basis of the sequence information obtained from the previous cloning step. In addition, the sequence of the entire region of LBVV RNA2 was confirmed by recloning the LBVV genomic RNAs using the LBVV-specific primers (the sequence of each primer is not shown) designed on the basis of the sequence information obtained by the genome-walking method.

---

**Fig. 1.** Schematic representation of LBVV RNA2 and the cloning strategies used in this work. Nucleotide positions of LBVV transcription initiation and termination on the viral RNA (V) are indicated by ▼ and △, respectively. The five major ORFs on the virus-complementary RNA (VC) are represented by boxes, and the nucleotide positions of the ORF start and end are indicated below the boxes. The six cDNA clones derived from the genome-walking method are shown as solid lines. The cDNA CP clone (Sasaya et al., 2001) was the starting point for the genome-walking method (hatched line). cDNA clones of both ends generated by 5′RACE and obtained by 3′RACE after polyadenylation of the LBVV genomic RNAs are shown as triple and dotted lines, respectively. cDNA clones resulting from direct RT-PCR using virus-specific primers in the second round of cloning are shown as arrows. The number of independent cDNA clones used for sequencing is shown in parentheses under the lines.
Table 1. Synthetic oligonucleotide primers used in this study

<table>
<thead>
<tr>
<th>Primer*</th>
<th>Nucleotide position</th>
<th>Primer sequence (3'→5')</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st GW</td>
<td>1135–1151</td>
<td>ATACCGTTGGCCTATCTTG</td>
<td>First-strand cDNA synthesis for genome walking</td>
</tr>
<tr>
<td>GW-1p</td>
<td>1342–1358</td>
<td>TAGAGTCCGACAGTGAAGAG</td>
<td>First genome walking</td>
</tr>
<tr>
<td>GW-2p</td>
<td>2543–2561</td>
<td>TGTTACATGCTGCTGCTCA</td>
<td>Second genome walking</td>
</tr>
<tr>
<td>GW-3p</td>
<td>3090–3108</td>
<td>CTGTTTAGCACCACAAATGCG</td>
<td>Third genome walking</td>
</tr>
<tr>
<td>GW-4p</td>
<td>4045–4063</td>
<td>AGCTGTTGGGAGGAGGAGAG</td>
<td>Fourth genome walking</td>
</tr>
<tr>
<td>GW-5p</td>
<td>5053–5071</td>
<td>CTTCTTATATCTCTGCGG</td>
<td>Fifth genome walking</td>
</tr>
<tr>
<td>GW-6p</td>
<td>5535–5553</td>
<td>GCTGTGATGAGTCTGCAGTCT</td>
<td>Sixth genome walking</td>
</tr>
<tr>
<td>p-5TA-R2-1p</td>
<td>505–519</td>
<td>CATCCACATCCGCTG</td>
<td>First cDNA synthesis for 5'RACE of virus-complementary RNA</td>
</tr>
<tr>
<td>5TA-R2-2p</td>
<td>384–401</td>
<td>GGGCTATTTGTGAGGTCC</td>
<td>First PCR for 5'RACE of virus-complementary RNA</td>
</tr>
<tr>
<td>5TA-R2-3p</td>
<td>351–369</td>
<td>TCACCTCAAGAACAGCATC</td>
<td>First PCR for 5'RACE of virus-complementary RNA</td>
</tr>
<tr>
<td>5TA-R2-1p</td>
<td>447–464</td>
<td>AGGTCTGTTGATCCCGTG</td>
<td>Second PCR for 5'RACE of virus-complementary RNA</td>
</tr>
<tr>
<td>5TA-R2-4p</td>
<td>289–306</td>
<td>TGCTTTATCATCCTCGG</td>
<td>Second PCR for 5'RACE of virus-complementary RNA and PCR for 3'RACE of viral RNA</td>
</tr>
<tr>
<td>p-3TA-R2-1p</td>
<td>5248–5262</td>
<td>CCATACCCGCGAGGAGAG</td>
<td>First-strand cDNA synthesis for 5'RACE of viral RNA</td>
</tr>
<tr>
<td>3TA-R2-1p</td>
<td>5282–5484</td>
<td>GGGCAGGCTGAAAGAGAGAGAG</td>
<td>First PCR for 5'RACE of viral RNA</td>
</tr>
<tr>
<td>LB3-6p</td>
<td>5654–5673</td>
<td>TGGGAGATATAGTGTTCAAG</td>
<td>First PCR for 5'RACE of viral RNA</td>
</tr>
<tr>
<td>3TA-R2-2p</td>
<td>5899–5916</td>
<td>CTGGAATGAAAGTGATGGGAAG</td>
<td>Second PCR for 5'RACE of viral RNA and PCR for 3'RACE of virus-complementary RNA</td>
</tr>
<tr>
<td>3TA-R2-3p</td>
<td>5535–5553</td>
<td>AGACGACACTCATCACAGC</td>
<td>Second PCR for 5'RACE of viral RNA</td>
</tr>
<tr>
<td>5'-PreEx*</td>
<td>1528–1549</td>
<td>GATGAGCTATTGCTGCTGCTG</td>
<td>Primer extension analysis for the 3' end of the ORF2 mRNA</td>
</tr>
<tr>
<td>RNase dig1p*</td>
<td>1407–1426</td>
<td>ACATATGATAGTAGGATCCC</td>
<td>RNase H analysis for the 3' end of the CP mRNA</td>
</tr>
<tr>
<td>RNase dig1p</td>
<td>1389–1408</td>
<td>CCATGAGAAAAGTTGAGGAG</td>
<td>RNase H analysis for the 3' end of the CP mRNA</td>
</tr>
<tr>
<td>RNase dig2p</td>
<td>1435–1454</td>
<td>TACCTATAGGCTTCTTCTTC</td>
<td>RNase H analysis for the 3' end of the CP mRNA</td>
</tr>
<tr>
<td>P1-N6</td>
<td>–</td>
<td>GCCCGGAGCTCTGCGAGAATTC</td>
<td>Second cDNA synthesis for genome walking</td>
</tr>
<tr>
<td>P2-T17</td>
<td>–</td>
<td>GGCCAGCCGCTGCGACTGACCT</td>
<td>First-strand cDNA synthesis for 3'RACE of viral RNA</td>
</tr>
<tr>
<td>5'-p-Adapter</td>
<td>–</td>
<td>CAATACCTTCTGACATGCTG</td>
<td>Adapter to the 3' end of mRNA</td>
</tr>
<tr>
<td>1st-3'-R</td>
<td>–</td>
<td>CATGCTGACTGACACTGCTAG</td>
<td>First-strand cDNA synthesis and first PCR for 3'RACE of mRNA</td>
</tr>
<tr>
<td>2nd-3'-R</td>
<td>–</td>
<td>TCGATGAGTCAGAAGTATTTG</td>
<td>Second PCR for 3'RACE of mRNA</td>
</tr>
</tbody>
</table>

*P, Positive-sense complementary to viral RNA; n, negative-, viral RNA-sense.

Since the purified LBVV RNA preparation contained viral and virus-complementary RNAs (Sasaya et al., 2001), the exact ends of the LBVV RNA2 were verified by a 5'RACE system using the 5'-Full RACE Core Set (TaKaRa Bio) as described previously (Sasaya et al., 2002). To determine the 3' end of the viral RNA of LBVV RNA2, LBVV genomic RNAs were reverse-transcribed using the 5'-end phosphorylated primer p-5TA-R2-1p. First-strand cDNA synthesis was followed by circularization of the 5' and 3' ends of the first-strand cDNA using RNA ligase (TaKaRa Bio). Subsequently, the ligated junction of the 5' and 3' ends was followed first by PCR amplification with virus-specific primers 5TA-R2-2p and 5TA-R2-3p and then by PCR amplification with virus-specific primers 5TA-R2-1p and 5TA-R2-4p. For the 5' end, first-strand cDNA was synthesized using the 5'-end phosphorylated primer p-3TA-R2-1p. Two pairs of virus-specific primers, 3TA-R2-1p with LB3-6p, and 3TA-R2-2p with 3TA-R2-3p, were used for first- and second-round PCR amplifications, respectively.

The sequences of both ends were also confirmed by 3'RACE after polyadenylation of the LBVV genomic RNAs as described previously (Sasaya et al., 2002). First-strand cDNA was synthesized from the 3'-polyadenylated LBVV RNAs using the first-strand cDNA synthesis primer P2-T17, which consisted of oligo(dT) linked to a known oligonucleotide at its 3' end. cDNA was amplified by PCR with the known oligonucleotides P2 and an LBVV-specific primer, either 5TA-R2-4p for the 3' end or 3TA-R2-2p for the 5' end of LBVV RNA2.

Determinação do 5' e 3' terminais de RNA LBVV. Para determinar os 5'-e 3'-terminais de sequência de nucleotídeos dos RNA LBVV, incluindo uma pequena ORF (designada gene 6) precedendo a polimerase L e a polimerase L no RNA1 (Sasaya et al., 2002), e para assegurar que os RNA LBVV moleculares tivessem um cap para o 5' end, cDNAs de ambos os correspodentes RNAs foram obtidos usando o primeiro escolhido RLM-RACE kit (Ambion), que especificamente seleciona transcritos de RNA com cinta intestinal fosfofrase (CIP) e ácido tabaco estrafofrase (TAP) tratamento (Liu & Gorovsky, 1993). Briefly, isolado poli(A)+ RNA de LBVV-infetados plantas de alface foi tratado com CIP para remover o 5' fosfato de todos os poli(A)+ RNA molecules except those that had a cap structure. TAP was then used to remove the cap structure from the poly(A)+ RNA, leaving a 5'-monophosphate. An RNA adapter oligonucleotide in the kit was ligated in the presence of T4 RNA ligase to the 5' end of the
CIP/TAP-treated poly(A)+ RNA. The ligated RNA was reverse-transcribed using a random primer, followed by first- and second-round PCR amplification with the 5’RACE outer primer in the kit and the first LBVV-specific 5’RACE mRNA primer (sequence not shown), which was designed on the basis of the LBVV genome sequence specific to the individual gene, and the 5’RACE inner primer in the kit and the second LBVV-specific 5’RACE mRNA primer.

To determine the nucleotide sequence at the 3’ ends of all seven LBVV mRNAs and to ascertain the actual size of the poly(A) tails, the sequence specific to the individual gene, and the 5’ and the first LBVV-specific 5’ of poly(A)+ RNA using T4 RNA ligase, as described by Liu & Gorovsky (1993). After excess oligonucleotide had been removed by centrifugal ultrafiltration, first-strand cDNA was synthesized using the first 3’RACE primer 5’-3’ R, whose sequence is complementary to that of 5’-p-Adapter. cDNA was amplified by PCR with 1 3’-R and the first LBVV-specific 3’ RACE mRNA primer (sequence not shown), which was designed on the basis of the LBVV genome sequence specific to the individual gene, followed by PCR amplification with the second 3’ RACE primer 2’3’-R and the LBVV-specific 3’RACE mRNA primer.

**Northern blot hybridization.** Poly(A)+ RNA samples from healthy and LBVV-infected lettuce plants were analysed by Northern blot hybridization as described previously (Kusaba et al., 1998). After heat denaturation in formaldehyde/formamide, poly(A)+ RNA was separated on 1.5% agarose horizontal submarine gels in MOPS/EDTA buffer containing 0.22 M formaldehyde and transferred to Hybond nitrocellulose membranes (Amersham-Pharmacia Biotech) by capillary blotting in 20 x SSC. Gel-purified inserts obtained from LBVV cDNA clones harbouring each LBVV gene were labelled with [α-32P]dCTP using a random primer labelling kit (Amersham-Pharmacia Biotech). Prehybridization, hybridization, washings and membrane exposures were carried out according to standard protocols (Sambrook & Russell, 2000).

**Primer extension analysis of the 5’ end of LBVV ORF2 mRNA.** The 5’ end of the ORF2 mRNA was mapped by primer extension analysis on poly(A)+ RNA from the LBVV-infected lettuce plants. Two micrograms of poly(A)+ RNA from healthy and LBVV-infected lettuce plants was reverse-transcribed with SuperScript II RNase H- Reverse Transcriptase (Invitrogen) using the LBVV-specific primer 5’-PrEx and [α-32P]dCTP. The radio-labelled cDNA was separated on an 8% polyacrylamide urea gel next to the products of a sequencing reaction of single-stranded cDNA from a plasmid harbouring the gene junction region between the CP and ORF2 using the same oligonucleotide primer.

**RNase H analysis of the 3’ end of LBVV CP mRNA.** The 3’ end of the CP mRNA was analysed by electrophoretic mobility shifts of Northern blot hybridizations following RNase H digestion, as described by Masters & Samuel (1984). Five micrograms of poly(A)+ RNA from LBVV-infected lettuce plants was incubated for 20 min at 25°C in a volume of 30 µl reaction mixture containing 20 mM Tris/HCl, pH 7.5, 10 mM MgCl2, 100 mM KCl and 0.1 mM DTT, with or without 1 µg of appropriate oligonucleotide primers. RNase H digestions were performed by the addition of 2 U RNase H (Gibco-BRL) and incubation for 30 min at 37°C. Digeste RNAs were recovered by phenol/chloroform extraction and ethanol precipitation prior to Northern blot hybridization using the LBVV CP-specific probe. The oligonucleotide primers used were as follows: RNase dig1, which annealed to the virus-complementary RNA from nt 1407 to 1426 of the LBVV RNA2; RNase dig2, which annealed to the viral RNA from nt 1389 to 1408; RNase dig2’, which annealed to the virus-complementary RNA from nt 1435 to 1454 and presumably does not exist in the CP mRNA; and oligo(dT)12-18, which annealed to the poly(A) sequence.

**RESULTS AND DISCUSSION**

**Coding sequence of LBVV RNA2**

Using the strategy described in Fig. 1, a total of six different PCR amplification steps were carried out to determine the entire nucleotide sequence of LBVV RNA2. LBVV RNA2 consisted of 6081 nt and contained antisense information for five major ORFs, which were designated ORF1 (nt 210–1403), ORF2 (nt 1493–2494), ORF3 (nt 2617–3489), ORF4 (nt 3843–4337) and ORF5 (nt 4530–5636) (Fig. 1). ORF1 corresponded to the previously described LBVV CP, which was distantly similar to the nucleocapsid protein of rhabdoviruses (Sasaya et al., 2001). The remaining four ORFs had coding capacities of 36, 32, 19 and 41 kDa, respectively.

The overall genomic organization of LBVV resembled that of plant rhabdoviruses (genera Nucleorhabdovirus and Cytorhabdovirus). Plant rhabdoviruses contain at least six genes encoding the nucleocapsid protein, phosphoprotein, non-structural protein, matrix protein, glycoprotein and L polymerase (Scholthof et al., 1994; Wetzel et al., 1994; Chen et al., 1998; Tanno et al., 2000). The matrix protein and glycoprotein form the viral envelope and are involved in the movement of virus into and out of host cells. The nucleocapsid protein, phosphoprotein and L polymerase form the viral nucleocapsids and are involved in expressing and replicating the virus genome (Conzelmann, 1998; Walker et al., 2000; Rose & Whitt, 2001; Barr et al., 2002). In the case of LBVV, its genome appeared to be split between the glycoprotein and L polymerase genes, with the L polymerase gene present on RNA1 and the other five genes present on RNA2 (see Fig. 3A). Although previous analysis of the LBVV CP and L polymerase revealed sequence similarity to the nucleocapsid proteins and the L polymerases of rhabdoviruses, respectively (Sasaya et al., 2001, 2002), FASTA and BLAST analyses of translated sequences from the four other LBVV genes showed no apparent similarity to other known viral sequences. However, this was not surprising since the primary sequences of the phospho-, non-structural, matrix and glycoproteins are poorly conserved among plant rhabdoviruses. The overall amino acid sequence identities of corresponding genes are less than 20% (Heaton et al., 1987; Hillman et al., 1990; Scholthof et al., 1994; Chen et al., 1998; Luo & Fang, 1998; Luo et al., 1998; Tanno et al., 2000).

**5’- and 3’-terminal sequences of the LBVV genome**

The 5’ and 3’ ends of the genomic RNA of rhabdoviruses are complementary and are thought to represent conserved promoters for RNA synthesis (Conzelmann, 1998). Complementarity between the extreme 5’ and 3’ ends of the LBVV RNA2 was low (Fig. 2A). Instead, nt 12–24 and 6047–6059 of LBVV RNA2 showed a 13 nt stretch of complementarity with a single mismatch, similar to that in LBVV RNA1 (Sasaya et al., 2002). A close examination
Transcription strategy of Lettuce big-vein virus

Fig. 2. Sequence analysis of LBVV genomic ends. (A) Complementarity between the 3′ and 5′ ends of LBVV RNA2. Vertical lines indicate complementary nucleotides between both extremities. (B) Comparison of the 3′ and 5′ ends of RNA1 and RNA2. The 3′- and 5′-terminal sequences for each RNA are shown in the viral RNA orientation. An asterisk indicates identical nucleotides in RNA1 and RNA2. The putative sequence (3′-ACUAAAA-5′) that is thought to operate as a promoter for replication is indicated by boxes.

Identification of LBVV mRNAs

Poly(A)^+ RNA from healthy and LBVV-infected lettuce plants was analysed by Northern hybridization with probes prepared from cDNA clones of all LBVV genes, including the L polymerase gene on RNA1 and a small ORF (designated gene 6) preceding it (Sasaya et al., 2002). As shown in Fig. 3, seven monocistronic transcripts of the expected sizes were detected with cDNA probes corresponding to each LBVV gene. None of the probes reacted with poly(A)^+ RNA prepared from healthy lettuce plants. The results clearly showed that each monocistronic mRNA of LBVV appeared to represent a full-length polyadenylated transcript of the corresponding gene in the LBVV-infected lettuce plants.

Analysis of LBVV mRNAs

We determined the 5′-terminal sequences of all LBVV mRNAs corresponding to the seven genes (Fig. 3A) by an RNA ligase-mediated RACE system, which specifically selects for capped transcripts and determines only their 5′ ends (Liu & Gorovsky, 1993). After nested PCR amplification, the sizes of the PCR fragments for the individual LBVV mRNAs were obtained, which suggested that LBVV mRNAs carry a cap structure on their 5′ ends. To obtain a representative set of sequences, 6–10 clones were randomly selected and their sequences analysed. Almost all of the cDNAs initiated with the sequence 5′-GAGA-3′, which was the exact complement of the corresponding LBVV genomic templates, 3′-CUCU-5′ (Fig. 4). In addition, the 5′ end of ORF2 mRNA was analysed by primer extension (Fig. 5). Autoradiography of the primer extension products using poly(A)^+ RNA from LBVV-infected lettuce plants revealed doublet bands corresponding to nt 1434 and 1435 of LBVV RNA2. No similar bands were detected in the primer extension products using poly(A)^+ RNA from healthy lettuce plants (Fig. 5). The more slowly migrating band (nt 1434) most likely resulted from copying of the cap structure at the 5′ end of ORF2 mRNA, as described in other viral mRNAs carrying a cap structure (Gupta & Kingsbury, 1984; Heaton et al., 1987; Zuidema et al., 1987; Hillman et al., 1990; Scholthof et al., 1994). The more rapidly migrating band (nt 1435) probably represented...
Fig. 3. Northern blot analysis of poly(A)⁺ RNAs from LBVV-infected lettuce plants. (A) Schematic representation of the genomic organization of LBVV and the location of probes. The six major LBVV ORFs and a small ORF upstream of the L polymerase gene are represented by boxes. V and VC indicate viral and virus-complementary RNAs, respectively. The probes representing each LBVV gene are indicated by a–g. (B) Northern blot analysis. Poly(A)⁺ RNAs from healthy or LBVV-infected lettuce plants are indicated by H and V, respectively. Each lane was hybridized with the probes indicated in (A). RNA molecular size markers (kb) (Perfect RNA Marker; Novagen) are shown on the left.

Fig. 4. Regulatory sequences of the LBVV genome. (A) LBVV genomic RNA sequences corresponding to the gene junctions. I, Sequences corresponding to the 3′ end of each mRNA; II, intergenic regions; III, sequences corresponding to the 5′ end of each mRNA. Transcription signals are shown in the viral RNA sense. (B) Consensus transcription termination/polyadenylation, intergenic region and initiation signals of four genera of rhabdoviruses, *Cytorhabdovirus*, *Nucleorhabdovirus*, *Vesiculovirus* and *Lyssavirus*. Consensus sequences were derived as described previously (Tordo et al., 1986; Heaton et al., 1989; Wetzel et al., 1994; Chen et al., 1998; Tanno et al., 2000). LNYV, Lettuce necrotic yellows virus; NCMV, Northern cereal mosaic virus; SYNV, Sonchus yellow net virus; RYSV, Rice yellow stunt virus; VSV, Vesicular stomatitis Indiana virus; RABV, Rabies virus. R, A or G; Y, C or U; W, A or U; N, number of variable nucleotides.
The 3′-terminal sequences of LBVV mRNAs were determined by cloning and sequencing of the 3′RACE-generated cDNA clones from poly(A) + RNA of LBVV-infected lettuce plants. To obtain the 3′-terminal sequences of these LBVV mRNAs, 6–10 clones derived from each mRNA were analysed. The lengths of their poly(A) tails differed, but all of their clones usually contained more than 100 A residues, with a maximum of 173 A residues. There were two types of poly(A) + RNA molecules by an approximately 100–200 nt compared with the untreated control. No such change was seen in the size of poly(A) + RNA molecules treated with RNase H in the presence of the virus-complementary RNA sense oligonucleotide primer RNase dig 1 p and viral RNA sense oligonucleotide primer RNase dig 2 p, which are located on gene 2. These results indicated that CP mRNA possessed a virus-complementary (positive-sense) RNA polarity and terminated before the initiation site of the ORF2 mRNA, which began at nt 1435 of LBVV RNA2, and that the poly(A) + RNA sequences were terminally located at the 3′ end of the CP mRNA.

**Sequences of gene junctions**

The gene junctions of rhabdoviruses consist of three parts: a gene-end sequence, which is required for transcription termination/polyadenylation; an intergenic region (IGR), which is not represented in the mRNAs; and a gene-start sequence, which plays an important role in transcription initiation. According to the stop-and-start model of transcription, the viral polymerase complex transcribes a gene, polyadenylates and releases the mRNA, and re-initiates transcription of the next gene under the direction of the gene junction sequences (Banerjee et al., 1991; Conzelmann, 1998; Barr et al., 2002). The determination of the 5′ and 3′ ends of all LBVV mRNAs described above led us to delineate precisely the gene junctions in the LBVV genome (Fig. 4).

A conserved sequence for the initiation of LBVV genes, 3′-CUCU-5′, was found. This conserved sequence is also found in the gene junction regions of LNYV and NCMV in the genus **Cytorhabdovirus** (Wetzel et al., 1994; Tanno et al., 2000). Although the transcription initiation consensus signal sequence of most rhabdoviruses, including genus **Nucleorhabdovirus**, is 3′-UUGU-5′ (Heaton et al., 1989; Walker et al., 2000), the 3′-CUCU-5′ sequence in the gene junction regions was considered to act as a transcription initiation signal in both LBVV and rhabdoviruses in the genus **Cytorhabdovirus**.

In all rhabdoviruses, gene-end sequences comprise an AU-rich region and a cytidylate and a poly(U) tract (Conzelmann, 1998). The poly(U) tract is thought to be a template for the mRNA poly(A) tail by reiterative
transcription or slippage (Barr et al., 2002). The gene-end sequences of LBVV were either 3′-UAUNCAUUUUUUU-5′ (type A) or 3′-UAUUUCUUUUUUU-5′ (type B), reminiscent of those observed in rhabdoviruses (Fig. 4). Interestingly, when the transcription termination/polyadenylation signal of LBVV was 3′-UAUNCAUUUUUUU-5′, the transcription of the next gene initiated immediately after the mononucleotide G (IGR), while for the type B sequence, the two genes (genes 5 and L) were located at the 5′-terminal ends of each RNA (viral RNA) or the transcription was re-initiated after a long 42 nt IGR (gene 4). These gene-end sequences of LBVV might play a role not only in the termination of transcription and polyadenylation but also in re-initiating transcription of the downstream genes. The role of gene-end sequences in signalling initiation of the downstream mRNA synthesis was recently reported for Vesicular stomatitis Indiana virus (VSIV) (Hinzman et al., 2002).

The IGRs of LBVV, except for the gene junction between genes 4 and 5, were highly conserved, consisting of the mononucleotide G (Fig. 4). The IGR of a fish rhabdovirus in the genus Novirhabdovirus also consists of a single nucleotide, G or A (Walker et al., 2000). The IGRs of rhabdoviruses either are highly conserved in sequence and length (genera Vesiculovirus and Nucleorhabdovirus) or are not conserved in either sequence or length (genera Lyssavirus, Ephemerovirus and Cytorhabdovirus), whereas the first residue of the IGRs in the most rhabdoviruses is G (Conzelmann, 1998; Walker et al., 2000). The G residue also followed immediately downstream of the last genes, gene 5 on RNA2 and the L polymerase gene on RNA1. The G residue might play an important role in the termination of LBVV mRNA transcripts, as reported for VSIV in which the first G residue in the IGR is required for efficient termination of the upstream transcript (Barr et al., 1997; Stillman & Whitt, 1997, 1998). The IGR between genes 4 and 5 differed from the other LBVV IGRs and consisted of 42 nt. Such a long IGR is also found in Rabies virus (RABV) (Tordo et al., 1986), Bovine ephemeral fever virus (Walker et al., 2000) and LNYV (Wetzel et al., 1994). The long IGR of RABV was reported to correlate with transcriptional attenuation (Finke et al., 2000). However, because of the varying specific activities of our probes, Northern blot analysis failed to show a different degree of transcription attenuation between genes 4 and 5 (Fig. 3).

Conclusion

This study completes the sequence analysis of the entire LBVV genome, which is the first report of the entire sequence of a member of the genus Varicosavirus. The LBVV genome is composed of 12,878 nt, divided into two segments. LBVV RNA1 consists of 6797 nt and contains a small gene 6 and L polymerase gene (Sasaya et al., 2002). LBVV RNA2 consists of 6081 nt and contains five genes. The genomic organization of LBVV and the amino acid sequences of LBVV CP and L polymerase were similar to those in rhabdovirus genomes. At the LBVV gene junctions, there were transcription termination/polyadenylation and initiation signals comparable with those of rhabdoviruses, and LBVV transcribed capped and polyadenylated monocistronic RNAs. LBVV and rhabdoviruses may utilize a similar mechanism to express individual genes differentially from a contiguous viral genome. These observations provide further evidence that LBVV is most closely related to viruses in the family Rhabdoviridae of the order Mononegavirales, even though LBVV has a genome that is divided into two segments and does not have enveloped virions. Furthermore, compared with viruses in the family Rhabdoviridae, purified LBVV contains a relatively large amount of positive-sense RNA (Sasaya et al., 2001), which is also a unique property of LBVV.

ACKNOWLEDGEMENTS

We are very grateful to Dr Mike A. Mayo for a critical reading of the manuscript and helpful advice, and to Mrs Yoshiko Aono for technical assistance. This research was funded by a grant from the Ministry of Agricultural, Forestry and Fisheries.

REFERENCES


T. Sasaya and others
Transcription strategy of Lettuce big-vein virus


