Virology 297, 289–297 (2002)
doi:10.1006/viro.2002.1420

The Nucleotide Sequence of RNA1 of *Lettuce big-vein virus*, Genus *Varicosavirus*, Reveals Its Relation to Nonsegmented Negative-Strand RNA Viruses

Takahide Sasaya,*[,1] Koichi Ishikawa,* and Hiroki Koganezawa†

*National Agricultural Research Center for Western Region, Zentsuji Campus, Zentsuji, Kagawa 765-8508, Japan; and †National Agricultural Research Center for Western Region, Fukuyama, Hiroshima 721-8514, Japan

Received December 12, 2001; accepted February 14, 2002

The complete nucleotide sequence of RNA1 from *Lettuce big-vein virus* (LBVV), the type member of the genus *Varicosavirus*, was determined. LBVV RNA1 consists of 6797 nucleotides and contains one large ORF that encodes a large (L) protein of 2040 amino acids with a predicted M_r of 232,092. Northern blot hybridization analysis indicated that the LBVV RNA1 is a negative-sense RNA. Database searches showed that the amino acid sequence of L protein is homologous to those of L polymerases of nonsegmented negative-strand RNA viruses. A cluster dendrogram derived from alignments of the LBVV L protein and the L polymerases indicated that the L protein is most closely related to the L polymerases of plant rhabdoviruses. Transcription termination/polyadenylation signal-like poly(U) tracts that resemble those in rhabdovirus and paramyxovirus RNAs were present upstream and downstream of the coding region. Although LBVV is related to rhabdoviruses, a key distinguishing feature is that the genome of LBVV is segmented. The results reemphasize the need to reconsider the taxonomic position of varicosaviruses. © 2002 Elsevier Science (USA)

Key Words: Lettuce big-vein virus; Varicosavirus; rhabdovirus; RNA polymerase; nonsegmented negative-strand RNA virus.

INTRODUCTION

Lettuce big-vein virus (LBVV) is the type member of the genus *Varicosavirus* and is currently classified as a double-strand RNA (dsRNA) virus (Mayo, 2000). The current members of the genus are LBVV, *Tobacco stunt virus* (TStV), *Camellia yellow mottle virus*, and *Freesia leaf necrosis virus*. Varicosaviruses are transmitted in soil by zoospores of a fungus vector and are retained stably in resting spores. Therefore, control of the disease is difficult once soil is infested with viruliferous fungi.

LBVV is thought to be an agent of big-vein disease in lettuce (*Lactuca sativa* L.), although a second virus, *Mirafiori lettuce virus*, was recently isolated from lettuce with big-vein symptoms (Roggero *et al.,* 2000). It occurs where winter to spring lettuce is grown in regions of Europe as well as in the United States, Australia, and Japan (Kuwata *et al.,* 1983; Vetten *et al.,* 1987; Huijberts *et al.,* 1990). LBVV virions are rod-shaped with modal lengths of 320–360 nm, have diameters of about 18 nm, and contain a coat protein (CP) with an M_r of 48,000 (Kuwata *et al.,* 1983). Previously, we demonstrated that LBVV virions do not contain segments of dsRNA but either of two segments of single-strand RNA (RNA1 and RNA2, formerly designated ss-1 and ss-2, respectively). The dsRNAs form in LBVV RNA preparations under different extraction conditions because positive-sense and negative-sense RNAs are separately encapsidated in the virions (Sasaya *et al.,* 2001). Moreover, database searches revealed that the CP of LBVV most resembled the nucleocapsid proteins of rhabdoviruses, nonsegmented negative-strand RNA viruses. Rhabdoviruses belong to the order *Mononegavirales* and have a linear nonsegmented negative-strand RNA genome and large enveloped virions with a prominent fringe of peplomers. In contrast, LBVV has a two-segmented negative-strand RNA genome and its particles are not enveloped.

To clarify the taxonomic position of genus *Varicosavirus*, we have cloned and sequenced the LBVV polymerase gene because RNA polymerase sequences are generally accepted as powerful criteria to use in determining taxonomic relationships (Poch *et al.,* 1990; Sidhu *et al.,* 1993). Judging from the segment size and the presence of the CP gene in LBVV RNA2, LBVV RNA1 was expected to harbor the RNA polymerase gene. In this article, we report the complete nucleotide sequence of the LBVV RNA1 and show that the LBVV RNA1 encodes one large protein (L protein) of 2040 amino acids. We also present evidence that the L protein of LBVV resembles RNA polymerases and most closely those of rhabdoviruses.

RESULTS

Purification of LBVV genomic RNAs and construction of a cDNA library of the LBVV RNA1

Preliminary attempts to construct a cDNA library of LBVV RNAs using RNA extracted from purified LBVV

[1] To whom correspondence and reprint requests should be addressed. Fax: +81-877-63-1683. E-mail: tsasaya@affrc.go.jp.

0042-6822/02 $35.00
© 2002 Elsevier Science (USA)
All rights reserved.

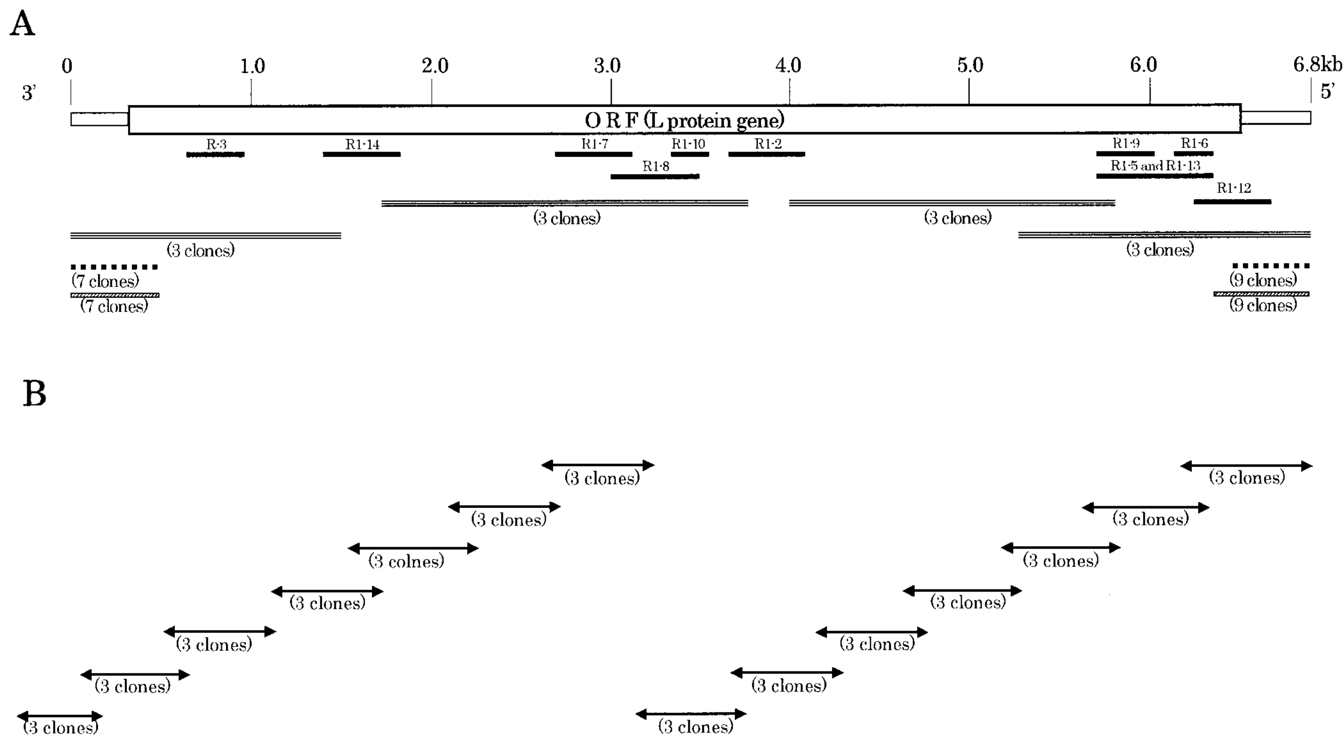

FIG. 1. A schematic representation of LBVV negative-sense RNA1 and the cloning strategies used in this work. The position of the ORF encoded by the complementary positive-sense RNA is indicated. (A) Eleven cDNA clones derived from a random-PCR method are shown as solid lines. cDNA clones synthesized by RT-PCR using the virus-specific primers are shown as triple lines. cDNA clones of both termini generated by 5′ RACE using 5′ Full RACE Core Set (Takara) and obtained by 3′ RACE after polyadenylation of the LBVV genomic RNAs are shown as dotted lines and shaded lines, respectively. (B) cDNA clones resulting from direct RT-PCR using the virus-specific primers in the second round of cloning. Parentheses under the lines show the number of independent cDNA clones used for sequencing.

particles by SDS–phenol extraction (Sasaya *et al.,* 2000) as a template were unsuccessful, probably because the purified LBVV genomic RNAs were not sufficiently abundant to allow the use of a standard cDNA library synthesis such as the Gubler–Hoffman method (Gubler and Hoffman, 1983). Virion purification from a large amount of infected lettuce leaves resulted in contamination of LBVV genomic RNAs with host DNAs and polysaccharides. The RNA preparations were further treated with RNaid reagents to eliminate these contaminants. LBVV RNA1 was purified by subsequent separation on a low melting temperature agarose gel. This method provided homogenous and relatively pure LBVV RNA1 that was used for the construction of a cDNA library of the LBVV RNA1 by a random PCR method (Froussard, 1992).

Sequencing strategies of the LBVV RNA1

The cloning strategies for determining the sequence of LBVV RNA1 are shown in Fig. 1. Using the random PCR method, a variety of LBVV RNA1 genome sequences were PCR-amplified and cloned. Twelve clones were randomly selected and sequenced. These 12 clones contained inserts of 134–491 nucleotides and among them two clones were identical. Only 1 of the 12 clones did not contain nucleotide sequence corresponding to that of the LBVV RNA1.

Two portions of the LBVV RNA1 were amplified by RT-PCR using oligonucleotide primers designed on the basis of the nucleotide sequences of the cDNA clones. The sequences of fragments 1911 and 2126 nucleotides in length were determined. To determine the sequence downstream of the established sequence, the LBVV RNAs were reverse-transcribed using a virus-specific primer designed on the basis of predetermined sequence. After tailing with poly(dC), the cDNA was amplified by PCR with primer AAP and a virus-specific primer. Seven clones from the dC-tailed PCR products were sequenced and six of seven clones included the sequence of the downstream primer. Two clones derived from the dC-tailed cDNA yielded 1513 nucleotides of sequence (excluding the poly(C) sequence) and started from the 3′ end of the LBVV RNA1, 3′-ACGCUCUGCGU···5′. The other four clones had deletions of 16, 48, 185, and 226 nucleotide residues from the 3′ end of the LBVV RNA1. In the same way, the sequence of the region upstream of the established sequence was determined by using a virus-specific primer for first-strand cDNA synthesis, and primer AAP with a virus-specific primer for PCR amplification. Nine clones from the dC-tailed PCR products were sequenced, two of which included the sequence of the upstream region. Two clones yielded 1471 nucleotides of sequence (excluding the

poly(C) sequence) and started two nucleotides downstream of the 5′ end of the LBVV RNA1.

Subsequently, the sequence of the entire 6.8-kb region was confirmed by recloning the LBVV genomic RNAs using virus-specific primers. A total of 14 independent cDNA cloning steps were carried out with the virus-specific primers and at least three clones were sequenced from each region (Fig. 1B).

Determination of the 3′ and 5′ ends of the LBVV RNA1

Since LBVV RNA preparations contained positive-sense and negative-sense RNAs (Sasaya *et al.,* 2001), both ends of the LBVV RNA1 were verified by a 5′RACE system and confirmed by 3′RACE on the 3′-polyadenylated LBVV RNAs. The sequence of the 3′ end of the LBVV RNA1 (negative sense) was determined by cloning and sequencing 5′ RACE-generated cDNA clones from the LBVV genomic RNAs using 5′ Full RACE Core Set (Takara). A total of seven clones were amplified from viral genomic RNAs as templates by PCR and their nucleotide sequences were determined. The clones derived from the 5′ RACE-generated cDNA showed variation in their 3′ end sequence: the sequence 3′-ACGCUCUGCGU· · ·5′ was found in four clones, 3′-CGCUCUGCGU· · ·5′ in two clones, and 3′-UCCUCUGCGU· · ·5′ in one clone. To confirm this result, purified LBVV genomic RNAs were polyadenylated, and the product was reverse-transcribed using primer AUAP-dT. Seven clones were amplified by PCR using AUAP and a virus-specific primer and their nucleotide sequences were determined. All seven clones that were sequenced had the same 3′-terminal sequence 3′-(A)nCGCUCUGCGU· · ·5′. From these results, the most likely 3′ end of the LBVV RNA obtained with two different methods was postulated to be 3′-ACGCUCUGCGU· · ·5′. Whether the nucleotide changes found at the 3′ end sequences of the clones derived from the 5′ RACE-generated cDNA were due to true variability in the LBVV RNA1 or were artifacts is not known.

The sequence of the 5′ end of the LBVV RNA1 (negative sense) was determined in the same way. A total of nine clones was amplified from viral genomic RNAs as templates by PCR and their nucleotide sequences were determined. The 5′ end sequence 3′-· · ·CAAACGCUGCGCC-5′ was found in all nine clones. The sequence of the 3′ end of the LBVV RNA1 was also confirmed by cloning and sequencing 3′ RACE-generated cDNA clones from the 3′-polyadenylated LBVV genomic RNAs. All nine clones obtained from the 3′ RACE-generated cDNA had the same 5′ terminal sequence 3′-· · ·CAAACGCUGCGCC(U)n-5′. These results suggested that the 5′ end of the LBVV RNA1 is 3′-· · ·CAAACGCUGCGCC-5′.

Using these strategies the complete nucleotide sequence of the LBVV RNA1 has been established. It consists of 6797 nucleotides and contains one large ORF (Fig. 1A).

Analysis of coding sequence of the LBVV RNA1

The single large ORF in the viral complementary RNA starts at an AUG codon at position 339–341 and ends at an UGA codon at position 6459–6461. The sequence around the initiation codon, CGAA<u>AUG</u>UC, corresponds only partially to the plant initiation consensus sequence AACA<u>AUG</u>GC (Lütcke *et al.,* 1987). The predicted protein encoded by the ORF comprised 2040 amino acids with an estimated molecular weight of 232,092. FASTA searches using the deduced amino acid sequence consistently retrieved the L proteins (polymerases) of non-segmented negative-strand RNA viruses within the order *Mononegavirales*. The core region of the predicted protein of LBVV was aligned with L polymerase of available plant rhabdoviruses and representatives of other genera in the order *Mononegavirales* (Fig. 2). The alignment revealed the conservation of functional domains of RNA-dependent RNA polymerases (Poch *et al.,* 1989, 1990; Müller *et al.,* 1994). The GHP (Gly-His-Pro) motif in block I, which is essential for polymerase activity of paramyxoviruses (Chandrika *et al.,* 1995), was present at position 348–350. The protein contained the conserved premotif A of block II, which presumably plays a role in RNA template binding and positioning (Müller *et al.,* 1994; Smallwood *et al.,* 1999). The protein also contained four motifs that correspond to the motifs A, B, C, and D of block III, thought to comprise the palm and finger regions of the polymerase active site (Poch *et al.,*1989, 1990 Müller *et al.,* 1994; Dhillon *et al.,* 2000). A GDN (Gly-Asp-Asn) at position 696–698 in motif C, and amino acids G (Gly) at position 662, and W (Trp) at position 671 in motif B, specific to mononegaviruses (Tordo *et al.,* 1992), were also conserved. In the conserved block VI, a motif $Kx_{17}GxGxG$ (Lys-x_{17}-Gly-x-Gly-x-Gly), proposed to be involved in ATP binding associated with polyadenylation or protein kinase activity (Poch *et al.,* 1990; Stec *et al.,* 1991; Mühlberger *et al.,* 1992; Dhillon *et al.,* 2000), was present at position 1643–1665 (Fig. 2). Furthermore, the small motifs LxSP (Leu-x-Ser-Pro) near the N-terminus as well as PHP (Pro-His-Pro), PYLGS (Pro-Tyr-Leu-Gly-Ser), HRF (His-Arg-Phe), and HxH (His-x-His) in block V, shown in previous alignments by Poch *et al.* (1990) and Sidhu *et al.* (1993), were located at positions 33–36, 1098–1100, 1134–1138, 1213–1215, and 1278–1280, respectively, although a few residues were substituted with chemically similar ones (data not shown). These results indicated that the large ORF on the LBVV RNA1 encodes the LBVV RNA polymerase and that the L protein of LBVV is related to the L proteins of mononegaviruses, even though LBVV is segmented. The LBVV RNA polymerase encoded on the LBVV RNA1 was designated L protein by analogy with proteins of other negative-strand RNA vi-

```
                 Block I                     Block II                       Block III
                 GHP motif                   Premotif A                     Motif A
LBVV   340 AFCLYRVWGHPTVD <164> KEREIKVAARMYSLMTERMRYYFVLTEGL <39> NINIDFSKWNTNMR <54>
NCMV   311 VMGLFRLWGHPEVD <166> KEREMNPVARMFALMTLKMRSYVVITENM <42> CINMDFEKWNLNMR <56>
SYNV   362 LHGLWRIWGHPIID <192> KEREMKTKARFFSLMSYKLRMYVTSTEEL <41> SMNIDFSKWNQNMR <54>
RYSV   355 LHGLYRIWAHPIID <168> KERELKIMARFFALLSFKMRLYFTATEEL <41> VINMDFVKWNQQMR <55>
VSIV   351 IYGSFRHWGHPFID <165> KERELKLAGRFFSLMSWKLREYFVITEYL <42> ANHIDYEKWNNHQR <56>
BEFV   377 FYSSFRHFGHPWID <165> KERELKEEGRFFSLMSYELRDYFVSTEYL <42> ANNIDYEKWNNYQR <56>
RABV   364 VYGCYRHWGHPYID <165> KERELKIEGRFFALMSWNLRLYFVITEKL <42> AFHLDYEKWNNHQR <58>
IHNV   316 FFGLFKHFAYPRVF <164> KEMELKIKGRGFGLMTFMPRLLQVLRESI <40> NKSLDINKFCTSQR <67>
MARV   371 LFSLQKHWGHPVLY <171> KEKELNI-GRTFGKLPYRVRNVQTLAEAL <47> SFVTDLEKYNLAFR <57>
ZEBOV  368 LFSIQKHWGHPVLH <171> KEKELNV-GRTFGKLPYPTRNVQTLCEAL <47> SFVTDLEKYNLAFR <57>
TRTV   357 MYFIFRIFGHPMVE <183> KERELSV-GRMFAMQPGKQRQVQILAEKL <49> SIVTDLSKFNQAFR <57>
HRSV   422 LYFLFRIFGHPMVD <183> KERELSV-GRMFAMQPGMFRQIQILAEKM <49> SIITDLSKFNQAFR <58>
MeV    349 IFSFFRSFGHPRLE <172> QEKEIKETGRLFAKMTYKMRACQVIAENL <95> FITTDLKKYCLNWR <57>
SeV    357 IFSFFRTFGHPSLE <172> KEKEIKQEGRLFAKMTYKMRAVQVLAETL <87> FLTTDLKKYCLNWR <57>
MuV    361 LLCIMRLWGHPTLT <178> KEKEIKATGRIFAKMTKRMRSCQVIAESL <83> FLTTDLTKYCLNWR <57>
BDV    184 VSSVQKSWYFPEIR <178> KEKELKVKGRFFSKQTLAIRIYQVVAEAA <38> VINLDYSSWCNGFR <54>

                                  B   l   o   c   k   III                              Block VI
                     Motif B                          Motif C              Motif D      Putative ATP binding motif
LBVV   654 MCYRGHLGGFEGLRQKGWTVATVCLL <12> LMGQGDNQII <64> LDGRQLPQWYKKT < 864> K <17> GDGTGG
NCMV   632 KSYEGHIRGFEGLRQKGWTVFTVVLI <12> LMGQGDNQVL <64> YKGVPLCSSLKRI < 913> K <24> GDGFGY
SYNV   705 WSRTGDESGKEGLRQKGWTITTVCDI <12> LIGGGDNQVL <63> YSGVPLRGRLKVI
RYSV   676 VCWIDDGAGKEGIRQKAWTIMTVCDI <12> LVGGGDNQVL <64> YKGVPLRSPLKQV
VSIV   671 VCWQGQEGGLEGLRQKGWTILNLLVI <12> VLAQGDNQVI <63> FRGVIRGLETKRW < 875> K <18> GDGSGG
BEFV   697 VCWEGQKGGLEGLRQKGWSILNYLMI <12> ILAQGDNQTI <63> IEGTIKGLPTKRW < 896> K <18> GDGSGG
RABV   686 TCWNGQDGGLEGLRQKGWSLVSLLMI <12> VLAQGDNQVL <63> FRGNILVPESKRW < 894> K <18> GDGSGG
IHNV   644 GVFSGLKGGIEGLCQYVWTICLLLRV <12> ILAQGDNVII <62> HCPQHLTLAIKKA < 854> K <22> GGGLGG
MARV   702 NAYHYHLGGIEGLQQKLWTCISCAQI <12> SSVMGDNQCI <60> LNGVQLPQSLKTM <1109> K <18> GEGSGA
ZEBOV  699 SSYRGHMGGIEGLQQKLWTSISCAQI <12> SAVMGDNQCI <60> LNGVQLPQSLKTA < 996> K <19> GEGAGA
TRTV   702 GLYRFHMGGIEGWCQKMWTMEAISLL <12> SLLNGDNQSI <59> SEGVMYPAAIKKV < 854> K <18> GEGAGN
HRSV   768 GLYRYHMGGIEGWCQKLWTIEAISLL <12> ALINGDNQSI <59> HNGVYYPASIKKV < 943> K <21> GEGAGN
MeV    730 IFIKYPMGGIEGYCQKLWTISTIPYL <12> SLVQGDNQTI <60> YDGLLVSQSLKSI < 915> K <21> GEGSGS
SeV    730 IFIHNPRGGIEGYCQKLWTLISISAI <12> AMVQGDNQAI <60> YDGKILPQCLKAL < 931> K <21> GEGAGA
MuV    736 IFIVSPRGGIEGLCQKLWTMISISTI <12> SMVQGDNQAI <60> YKGRILTQALKNV < 935> K <21> AEGSGA
BDV    511 TCAVGTKTMGEGMRQKLWTILTSCWE <12> ILGQGDNQTI <51> FRGVPVPGCLKQL
```

FIG. 2. Alignment of conserved motifs of L protein of LBVV, seven rhabdoviruses (NCMV, SYNV, RYSV, VSIV, BEFV, RABV, and IHNV), two filoviruses (MARV and ZEBOV), five paramyxoviruses (TRTV, HRSV, MeV, SeV, and MuV), and one bornavirus (BDV) within the order *Mononegavirales*. Numbers at the beginning of the lines indicate the position of the first displayed amino acid. Numbers within the brackets indicate the numbers of amino acids not represented in the figure. Conserved residues recognized previously for nonsegmented negative-strand RNA virus polymerases (Tordo *et al.,* 1992; Mühlberger *et al.,* 1992; Müller *et al.,* 1994; Dhillon *et al.,* 2000) are shown in bold letters. Additional strictly conserved residues are underlined. The abbreviated names of virus species and GenBank accession numbers used for analysis are described under Materials and Methods.

ruses. ORF analysis of the complementary strand showed several short ORFs, each with a coding capacity of less than 15kDa, seeming to be not functional.

Relationship of the LBVV L protein to the RNA polymerases of negative-strand RNA viruses

The predicted amino acid sequence of the LBVV L protein was compared to that of mononegaviruses by pairwise and multiple alignments. In the pairwise comparisons with RNA of *Northern cereal mosaic virus* (NCMV), *Sonchus yellow net virus* (SYNV), or *Rice yellow stunt virus* (RYSV), the overall identity of the LBVV L protein was 25.9% to NCMV, 25.0% to SYNV, and 24.3% to RYSV. This is lower than the overall sequence relatedness between the SYNV and RYSV L proteins, which is 27.8%, but is higher than the relatedness between the LBVV L protein and those of *Vesicular stomatitis Indiana virus* (VSIV) (24.1%), *Rabies virus* (RABV) (23.4%), *Bovine ephemeral fever virus* (BEFV) (23.4%), and *Infectious hematopoietic necrosis virus* (IHNV) (21.0%). The identity of the LBVV L protein with those of rhabdoviruses is also higher than the relatedness between the LBVV L protein and those of viruses in the *Paramyxoviridae*, *Filoviridae*, and *Bornaviridae*. The cluster dendrogram in Fig. 3, which is based on the multiple alignment illustrated in Fig. 2, shows that the putative LBVV polymerase clusters with the L proteins of plant rhabdoviruses and indicates that the L protein of LBVV is more distantly related to the L proteins of plant rhabdoviruses (genera *Cytorhabdovirus* and *Nucleorhabdovirus*) than they are to each other, but is more closely related to those of plant rhabdoviruses than to those of other rhabdoviruses.

Analysis of the 3′- and 5′-noncoding regions

The 3′-noncoding sequence of the RNA1 of LBVV is 338 nucleotides in length, extending from the first nucleotide to the translation initiation codon for the putative L protein. The 3′ end sequence of RNA1 had a high A + U content, 62.3%, similar to that of the 3′ leader sequences of mononegaviruses. Complementarity between the 3′ and 5′ end sequences is a common feature of the genomes of negative-strand RNA viruses. However, the complementarity between the extreme 3′ and 5′ ends of the LBVV RNA 1 is low (Fig. 4). Instead, base pairing between positions 24–43 and 6754–6773 in the viral negative-sense RNA can potentially give rise to a

FIG. 3. Phylogenetic relationship between LBVV and seven rhabdoviruses shown in Fig. 2. The phylogenetic tree was constructed based upon a CLUSTAL W multiple alignment of conserved region between the GHP motif and the motif D of L polymerase in Fig. 2 using the BLOSUM weighting matrix. The trees were generated using the neighbor-joining method and the MEGA software. The figures on the branches represent the percentage of trees containing each cluster out of 1000 bootstrap replicates. The branch lengths are proportional to the genetic distances between sequences. The scale bar indicates substitutions per amino acid site.

panhandle structure for the genome RNA since these regions showed a 20 nucleotide stretch of complementarity with four mismatches. Downstream from the large ORF, a stretch of six uridine residues (poly(U)-tract), which resemble transcription termination/polyadenylation signals of mononegaviruses (Heaton *et al.,* 1989; Sidhu *et al.,* 1993; Dhillon *et al.,* 2000), was identified at positions 6541–6546. A similar poly(U)-tract was also identified at position 275–281 in the 3′-noncoding region of LBVV RNA1.

Negative-sense nature of RNA1

To confirm the location of the LBVV L protein gene, LBVV RNA preparations were heat-denatured, annealed at room temperature, and then analyzed by Northern blot hybridization using positive-sense and negative-sense riboprobes of the 3′-terminal region of the LBVV L protein gene (Fig. 5). A positive-sense riboprobe hybridized with both single- and double-stranded RNA1; however, a negative-sense riboprobe hybridized only with double-stranded RNA1. These results indicate that the L protein gene of LBVV is located on the LBVV RNA1 and that negative-sense RNAs are more abundant than positive-sense RNAs, corresponding with our previous results (Sasaya *et al.,* 2001).

DISCUSSION

LBVV and TStV are serologically related viruses that have been classified in a new genus, *Varicosavirus* (Mayo, 2000). Specific defining characteristics of these viruses, including nucleotide sequence data, are lacking. The very small amounts of virion RNA available has impeded progress in obtaining sequences. However, a random PCR method developed by Froussard (1992) made it possible to construct an LBVV cDNA library from minute amounts of LBVV RNA1. This has led to the rapid production of the first complete segment sequence of a varicosavirus.

In the putative 3′-noncoding region of the LBVV RNA1, one small ORF is located between positions 108 and 230 of the viral negative-sense RNA and has the potential to encode a polypeptide of only 41 amino acids. The polypeptide is relatively hydrophobic. Analysis of the nucleotide sequence of NCMV (genus *Cytorhabdovirus*) (Tanno *et al.,* 2000) showed it to contain in the same position (immediately upstream from the L protein gene) a small ORF that encodes a hydrophobic polypeptide of 52 amino acids. A similar small ORF is also located between the G and L protein genes of RYSV (genus *Nucleorhabdovirus*) (Fang *et al.,* 1994). The conserved sequences of transcription termination/polyadenylation signals of mononegaviruses were identified after the UGA stop codons of these small ORFs. However, in preliminary experiments we were unable to detect an mRNA for the LBVV polypeptide by Northern blotting (data not shown) and no polypeptide was detected by Western blotting using an antiserum raised against a synthetic peptide from the 41 amino acid polypeptide (data not shown). However, more detailed experiments are needed to be sure that the small ORF in LBVV RNA1

```
3' ACGCUCUGCGUAUAAGUAUGAUGUUUAGGUUGUUUAGUUUUGGAAAAUUGCGGGCUAUUUGUCGGAAAGCAGCUCCGUAAGUCG····
5'CCGCGUCGCAAACACAGCACAACAAAAUCAUAAAAAUCAAGACCCCGACUAACACGGCUUUGUUAGAACUCAACUUGCUUUAGCG····
```

FIG. 4. Complementarity between the 3′ and 5′ ends of the LBVV RNA1. Sequences are shown as viral negative-sense RNA.

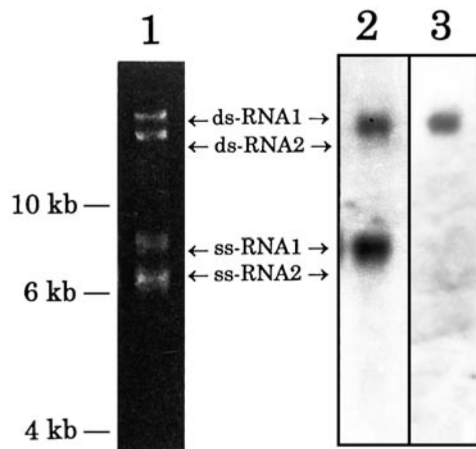

FIG. 5. Northern blot hybridization analysis of LBVV RNAs with riboprobes of the partial LBVV L protein gene. Lane 1 contained LBVV RNAs, which were stained with ethidium bromide after heat denaturation followed by annealing at room temperature. Lanes 2 and 3 contained LBVV RNAs hybridized with positive- and negative-sense riboprobes, respectively. The positions of single-strand RNA size markers (kb) are indicated on the left.

is not expressed because ORFs that encode small hydrophobic integral membrane proteins have been reported for viruses in the mononegavirus genera *Pneumovirus* (Perez *et al.,* 1997), *Rubulavirus* (Takeuchi *et al.,* 1996; He *et al.,* 1998), *Metapneumovirus* (Ling *et al.,* 1992), and *Ephemerovirus* (Wang *et al.,* 1994; McWilliam *et al.,* 1997).

The catalytic subunit of the polymerase complex (L protein) is the only region that is sufficiently conserved to allow the establishment of evolutionary relationships between negative-strand RNA viruses (Poch *et al.,* 1990). Even though LBVV is segmented, most of the conserved motifs previously identified in the L proteins of mononegaviruses (Poch *et al.,* 1989, 1990; Choi *et al.,* 1992; Tordo *et al.,* 1992) are also conserved in the LBVV L protein. In particular, the conserved sequence GDN (Gly-Asp-Asn) in motif C, which is likely to constitute the evolutionary and functional equivalent of the GDD (Gly-Asp-Asp) motif in the polymerases of positive-strand RNA viruses (Poch *et al.,* 1998, 1990; Choi *et al.,* 1992; Tordo *et al.,* 1992), is present. This is in striking contrast to segmented negative-strand RNA viruses whose polymerase has an SDD (Ser-Asp-Asp) motif. Furthermore, the tetrapeptide E(F/Y)xS (Glu(Phe/Try)xSer) located downstream from motif D, which was described as motif E and is specifically conserved in the polymerases of segmented negative-strand RNA viruses (Müller *et al.,* 1994), was not found in the L protein of LBVV. Phylogenetic data and pairwise amino acid sequence comparisons of the L protein between LBVV and those of mononegaviruses also suggest that the L protein of LBVV is more closely related to those of *Rhabdoviridae* viruses than to those of *Bornaviridae*, *Filoviridae*, and *Paramyxoviridae* viruses.

As well as the similarities in amino acid sequences of the nucleocapsid proteins (Sasaya *et al.,* 2001) and the L proteins, LBVV resembles rhabdoviruses in possessing conserved transcription termination/polyadenylation signal-like poly(U) tracts downstream and upstream of the L protein gene. In mononegaviruses, the conservative sequences surrounding the poly(U) tract act as signals for transcription termination/initiation, and the poly(U) tract is a template for polyadenylation of the nascent mRNA by a stuttering mechanism (Conzelmann, 1998). The sequence UAUGCA at positions 269–274 in the viral negative-sense RNA, immediately downstream of a poly(U) tract, is similar to the sequence UAUACA at the end of the LBVV CP gene on the LBVV RNA2 (Sasaya *et al.,* 2001). The presence of poly(U) tracts in the noncoding regions of LBVV RNA1 and RNA2 suggests that LBVV transcription is regulated by a mechanism similar to that which regulates gene transcription of mononegaviruses. However, mRNA for L protein was not detected by Northern blotting using the LBVV L protein gene-specific probe (data not shown), perhaps because insufficient RNA was present. This may have been because LBVV particles were present only at low concentration in infected lettuce plants; viral mRNAs accumulate in the infected lettuce plants only at limited periods during multiplication, and/or the L protein mRNA is the least abundant viral transcript.

LBVV is not the only plant virus that has a polymerase which resembles those of mononegaviruses but that has a bipartite genome. The genome organization and morphology of orchid fleck virus (OFV) are similar to those of rhabdoviruses, although the genome of OFV is divided into two segments (Calisher *et al.,* 2000). In the Sixth Report of the International Committee on Taxonomy of Viruses (ICTV), OFV was classified as an unassigned plant rhabdovirus (Wunner *et al.,* 1995), but it is currently classified as unassigned to any existing genus or family because of the segmentation of its genome (Calisher *et al.,* 2000). LBVV should possibly be classified in the order *Mononegavirales* together with OFV, but an obstacle to this classification is the derivation of the name *Mononegavirales* (single component negative-sense (RNA) virus). In conclusion, the results described here are the first report of the sequence of the polymerase gene of a member of genus *Varicosavirus* and show a molecular relationship between LBVV (genus *Varicosavirus*) and rhabdoviruses. The taxonomic position of varicosaviruses, which is currently classified as a dsRNA virus by the ICTV, may need to be reevaluated in light of our results.

MATERIALS AND METHODS

Virus and RNA

LBVV was purified from leaves showing typical big-vein symptoms as previously described (Sasaya *et al.,*

2001). The LBVV genomic RNAs were extracted from purified virus particles by SDS–phenol treatment followed by ethanol precipitation (Sasaya *et al.,* 2001). The LBVV RNAs were further treated with RNaid reagents (Bio 101, Inc.) to eliminate contaminating host DNAs and polysaccharides, according to the protocol provided by the supplier. The LBVV RNA1 was purified by subsequent separation on a low melting temperature agarose gel (Sambrook and Russell, 2000) and then used as a template for cDNA synthesis.

Cloning and sequencing of the LBVV RNA1

The cloning strategies for LBVV RNA1 are shown in Fig. 1. The PCR-based cDNA library of LBVV RNA1 was constructed as described by Froussard (1992). Briefly, first-strand cDNA was synthesized using a 26 nucleotide primer containing a random hexamer at its 3′-end (5′-GCCGGAGCTCTGCAGAATTCNNNNNN-3′) and SuperScript II RNase H^- Reverse Transcriptase (Gibco BRL) following the manufacturer's instructions. The second-strand cDNA was synthesized using Klenow fragment (TOYOBO), followed by PCR amplification with the universal primer (5′-GCCGGAGCTCTGCAGAATTC-3′).

To determine the nucleotide sequence between two clones, R1-14 and R1-2, two oligonucleotide primers, 5R1-1 (5′-AATGTTGAGATGGATGGAGG-3′; complementary to the positions 1784–1803, positive sense) and 3R1-1 (5′-TCTGATCTTTGGCATTGGTC-3′; complementary to the positions 3675–3694, negative sense), were designed on the basis of the sequences of these two clones. In the same way, the nucleotide sequence between R1-2 and R1-9 was determined using two primers, 5R1-2 (5′-ATCCCTCAGACAACCTA-3′; positions 3940–3956, positive sense) and 3R1-2 (5′-GTAAGAGGATGAGGCAAAAT-3′; positions 6046–6065, negative sense).

The cDNA libraries including the upstream or downstream regions were constructed from the LBVV RNA template. To determine the nucleotide sequence of the upstream region, a virus-specific primer, 5RA-R1-1 (5′-GTACCTATGATGCTTCTTGA-3′; positions 1513–1532, negative sense), was used for first-strand cDNA synthesis. After first-strand cDNA synthesis, the excess primer was removed using a GlassMAX DNA isolation spin cartridge system (Gibco BRL). Prior to PCR amplification, the cDNA was tailed with poly(dC) by terminal deoxynucleotidyl transferase (Gibco BRL). Primer AAP (5′-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3′) and a virus-specific primer, 5RA-R1-2 (5′-AAGAAGGTTGACATGAA-3′; positions 1497–1513, negative sense), were used to amplify the poly(dC)-tailed cDNA. The cDNA libraries including the downstream region of the LBVV RNA1 were constructed in the same way, since the LBVV RNA preparation contained positive- and negative-sense RNAs (Sasaya *et al.,* 2001). A virus-specific primer 5R1-5-4 (5′-TGTAGTAGTCATGGGAGAT-3′; positions 5306–5324, positive sense) was used for first-strand cDNA synthesis. Primer AAP and a virus-specific primer 5R1-5-5 (5′-TGGCACTGGAGGTACA-3′; positions 5324–5339, positive sense) were used to amplify the poly(dC)-tailed cDNA.

RACE for determination of both termini of the LBVV RNA1

Since the LBVV RNAs preparation contained positive- and negative-sense RNAs (Sasaya *et al.,* 2001), both ends of the LBVV RNA1 were verified by a 5′-RACE system; 5′ Full RACE Core Set (Takara) according to the protocols provided by suppliers. Both ends of the LBVV RNA1 were also confirmed by 3′ RACE (Frohman *et al.,* 1988) on the 3′-polyadenylated LBVV RNAs.

To determine the 3′ end of the LBVV RNA1 (negative sense), purified viral genomic RNAs were used as a template. First-strand cDNA was synthesized using the 5′ end-phosphorylated primer pgTA-R1-5 (5′-CAATACCTTCTGACC-3′; complementary to positions 820–834, negative sense), and cDNA was then circularized using RNA ligase (Takara). This was followed first by PCR amplification with virus-specific primers 5R1-3 (5′-ATAGAGATCGGGTTTCAGA-3′; positions 570–588, positive sense) and 5RA-R1-4 (5′-TTGACCTTGTAAACGGGGG-3′; positions 487–505, negative sense), and then by PCR amplification with virus-specific primers 5R1-3-1 (5′-GCTTCAACTGCTTTCTCAA; positions 772–790, positive sense) and 5RA-R1-6 (5-ATCCCATCCGTAATCCG-3′; positions 394–410, negative sense). In the same way, to determine the exact 5′ end of the LBVV RNA1 (negative sense), first-strand cDNA was synthesized by using the 5′ end-phosphorylated primer pgTA-R1-3 (5′-GAAGATTGCCACCTG-3′; positions 6386–6400, positive sense). Two pairs of virus-specific primers, 3TA-R1-1 (5′-AAACGAACCCAGATAGCTC-3′; positions 6544–6562, positive sense) with 3R1-4-1 (5′-ACATCATTCATCCTCCAAG-3′; positions 6467–6485, negative sense), and 3TA-R1-2 (5′-GAAAATATCCGAGCGTGTG-3′; position 6632–6650, positive sense) with 3R1-4 (5′-TATCCATCCACCGCTTGC-3′; position 6423–6440, negative sense), were used for first and second PCR amplifications, respectively.

To confirm the sequences obtained for the ends of RNA1, LBVV RNAs were treated with DNase I and 3′ polyadenylated (Ashulin *et al.,* 1992). First-strand cDNA was synthesized using an oligonucleotide primer (AUAP-dT) that consisted of a polylinker and oligo(dT) [5′-GGCCACGCGTCGACTAGTAC$(T)_{17}$-3′]. After excess primer had been removed using the GlassMAX DNA isolation spin cartridge system, cDNA was amplified by PCR with primer AUAP (5′-GGCCACGCGTCGACTAGTAC-3′) and a virus-specific primer, either 5RA-R1-6 for the 3′ end or 3TA-R1-1 for the 5′ end of the LBVV RNA1.

Cloning of PCR product and sequencing

The PCR protocol consisted of 3 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 56°C, and 2 min at 72°C, and finally an extension time of 10 min at 72°C. Amplified PCR fragments were separated by electrophoresis in a 1% agarose gel and isolated from the gel using a QIAquick Gel Extraction Kit (Qiagen). Purified PCR fragments were ligated into plasmid pGEM-T Easy Vector (Promega). The sequences of both strands were determined by the dideoxynucleotide chain termination method according to standard protocols (ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit, Perkin–Elmer) using an automated ABI Prism 377 DNA sequencer.

Analysis of nucleotide and amino acid sequences

The GENETIX-WIN version 4.3 (Software Development) program was used for analysis of nucleotide and amino acid sequence data. The amino acid sequence of the L protein of LBVV was compared with the DDBJ database using the TFASTA program (Pearson and Lipman, 1988). The amino acid sequence of the L protein of LBVV was compared with the L protein of mononegaviruses. Abbreviations and GenBank accession numbers of viruses used for the comparison of amino acid sequences are as follows: *Borna disease virus* (BDV), U04608; *Marburg virus* (MARV), Z12132; *Zaire Ebola virus* (ZEBOV), AF272001; *Measles virus* (MeV), K01711; *Sendai virus* (SeV), AB039658; *Mumps virus* (MuV), AB040874; *Human respiratory syncytial virus* (HRSV), AF013254; *Turkey rhinotracheitis virus* (TRTV), U65312; *Vesicular stomatitis Indiana virus* (VSIV), J02428; *Rabies virus* (RABV), M13215; *Bovine ephemeral fever virus* (BEFV), AF234533; *Infectious hematopoietic necrosis virus* (IHNV), L40883; *Northern cereal mosaic virus* (NCMV), AB030277; *Sonchus yellow net virus* (SYNV), L32603; and *Rice yellow stunt virus* (RYSV), AB011257. Alignment of the L proteins was generated by the program CLUSTAL W (Tompson *et al.,* 1994). A cluster dendrogram was constructed based on a CLUSTAL W multiple alignment of the conserved region between the GHP motif and motif D of L polymerase (Fig. 2) using the BLOSUM weighting matrix. The cluster dendrogram was generated by the neighbor-joining (NJ) method (Saitou and Nei, 1987), and the pairwise distances between the sequences were calculated using the MEGA version 2 software.

Northern blot analysis

The LBVV genomic RNAs were extracted from purified virus particles by SDS–phenol treatment. RNA was heat-denatured, annealed at room temperature, and then electrophoresed through a 1% nondenaturing agarose gel as described by Sasaya *et al.* (2001). LBVV RNAs were transferred onto a Hybond-N+ membrane (Amersham-Pharmacia) using 20× SSC as the transfer buffer. A 1.9-kb fragment of the partial LBVV polymerase gene, which was amplified by RT-PCR using virus-specific primers, 5R1-1 and 3R1-1, was ligated into plasmid pGEM-T Easy. Riboprobes were prepared from the plasmid harboring the partial LBVV polymerase gene using an RNA Labeling Kit (Amersham-Pharmacia). Prehybridization, hybridization, washings, and membrane exposures were carried out according to standard protocols (Sambrook and Russell, 2000).

ACKNOWLEDGMENTS

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

Sequence data from this article have been deposited with the EMBL/GenBank/DDBJ Data Libraries under Accession No. AB075039.

REFERENCES

Ashulin, L., Mawassi, M., and Bar-Joseph, M. (1992). Procedure to amplify cDNA from viroid RNA templates using the polymerase chain reaction. *Methods Mol. Cell. Biol.* **3,** 83–89.

Calisher, C. H., Carstens, E. B., Christian, P., Mahy, B. W. J., Mayo, M. A., and Shope, R. E. (2000). Unassigned viruses. *In* "Virus Taxonomy, Seventh Report of the International Committee on Taxonomy of Viruses" (M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner, Eds.), pp. 995–1008. Academic Press, San Diego.

Chandrika, R., Horikami, S. M., Smallwood, S., and Moyer, S. A. (1995). Mutations in conserved domain I of the Sendai virus L polymerase protein uncouple transcription and replication. *Virology* **213,** 352–363, doi:10.1006/viro.1995.0008.

Choi, T.-J., Kuwata, S., Koonin, E. V., Heaton, L. A., and Jackson, A. O. (1992). Structure of the L (polymerase) protein gene of sonchus yellow net virus. *Virology* **189,** 31–39.

Conzelmann, K.-K. (1998). Nonsegmented negative-strand RNA viruses: Genetics and manipulation of viral genomes. *Annu. Rev. Genet.* **32,** 123–162.

Dhillon, J., Cowley, J. A., Wang, Y., and Walker, P. J. (2000). RNA polymerase (L) gene and genome terminal sequences of ephemeroviruses bovine ephemeral fever virus and Adelaide River virus indicate a close relationship to vesiculoviruses. *Virus Res.* **70,** 87–95.

Fang, R.-X., Wang, Q., Xu, B.-Y., Pang, Z., Zhu, H.-T., Mang, K.-Q., Gao, D.-M., Qin, W.-S., and Chua, N.-H. (1994). Structure of the nucleocapsid protein gene of rice yellow stunt rhabdovirus. *Virology* **204,** 367–375, doi:10.1006/viro.1994.1541.

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.

Froussard, P. (1992). A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Res.* **20,** 2900.

Gubler, U., and Hoffman, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25,** 263–269.

He, B., Leser, G. P., Paterson, R. G., and Lamb, R. A. (1998). The paramyxovirus SV5 small hydrophobic (SH) protein is not essential for virus growth in tissue culture cells. *Virology* **250,** 30–40, doi:10.1006/viro.1998.9354.

Heaton, L. A., Hillman, B. I., Hunter, B. G., Zuidema, D., and Jackson, A. O. (1989). Physical map of the genome of sonchus yellow net virus,

a plant rhabdovirus with six genes and conserved gene junction sequences. *Proc. Natl. Acad. Sci. USA* **86,** 8665–8668.

Huijberts, N., Blystad, D.-R., and Bos, L. (1990). Lettuce big-vein virus: Mechanical transmission and relationships to tobacco stunt virus. *Ann. Appl. Biol.* **116,** 463–475.

Kuwata, S., Kubo, S., Yamashita, S., and Doi, Y. (1983). Rod-shaped particles, a probable entity of lettuce big vein virus. *Ann. Phytopathol. Soc. Jpn.* **49,** 246–251.

Ling, R., Easton, A. J., and Pringle, C. R. (1992). Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus and their intergenic regions reveals a gene order different from that of other pneumoviruses. *J. Gen. Virol.* **73,** 1709–1715.

Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F., and Scheele, G. A. (1987). Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6,** 43–48.

Mayo, M. A. (2000). Genus Varicosavirus. *In* "Virus Taxonomy, Seventh Report of the International Committee on Taxonomy of Viruses" (M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner, Eds.), pp. 521–523. Academic Press, San Diego.

McWilliam, S. M., Kongsuwan, K., Cowley, J. A., Byrne, K. A., and Walker, P. J. (1997). Genome organization and transcription strategy in the complex G_{NS}-L intergenic region of bovine ephemeral fever rhabdovirus. *J. Gen. Virol.* **78,** 1309–1317.

Mühlberger, E., Sanchez, A., Randolf, A., Will, C., Kiley, M. P., Klenk, H.-D., and Feldmann, H. (1992). The nucleotide sequence of the L gene of Marburg virus, a filovirus: Homologies with paramyxoviruses and rhabdoviruses. *Virology* **187,** 534–547.

Müller, R., Poch, O., Delarue, M., Bishop, D. H. L., and Bouloy, M. (1994). Rift valley fever virus L segment: Correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J. Gen. Virol.* **75,** 1345–1352.

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.

Perez, M., García-Barreno, B., Melero, J. A., Carrasco, L., and Guinea, R. (1997). Membrane permeability changes induced in *Escherichia coli* by the SH protein of human respiratory syncytial virus. *Virology* **235,** 342–351, doi:10.1006/viro.1997.8696.

Poch, O., Blumberg, B. M., Bougueleret, L., and Tordo, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: Theoretical assignment of functional domains. *J. Gen. Virol.* **71,** 1153–1162.

Poch, O., Sauvaget, I., Delarue, M., and Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **8,** 3867–3874.

Roggero, P., Ciuffo, M., Vaira, A. M., Accotto, G. P., Masenga, V., and Milne, R. G. (2000). An *Ophiovirus* isolated from lettuce with big-vein symptoms. *Arch. Virol.* **145,** 2629–2642.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425.

Sambrook, J., and Russell, D. W. (2000). "Molecular Cloning: A Laboratory Manual," 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sasaya, T., Ishikawa, K., and Koganezawa, H. (2001). Nucleotide sequence of the coat protein gene of *Lettuce big-vein virus. J. Gen. Virol.* **82,** 1509–1515.

Sidhu, M. S., Menonna, J. P., Cook, S. D., Dowling, P. C., and Udem, S. A. (1993). Canine distemper virus L gene: Sequence and comparison with related viruses. *Virology* **193,** 50–65, doi: 10.1006/viro.1993.1102.

Smallwood, S., Easson, C. D., Feller, J. A., Horikami, S. M., and Moyer, S. A. (1999). Mutations in conserved domain II of the large (L) subunit of the Sendai virus RNA polymerase abolish RNA synthesis. *Virology* **262,** 375–383, doi:10.1006/viro.1999.9933.

Stec, D. S., Hill, M. G., III, and Collins, P. L. (1991). Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* **183,** 273–287.

Takeuchi, K., Tanabayashi, K., Hishiyama, M., and Yamada, A. (1996). The mumps virus SH protein is a membrane protein and not essential for virus growth. *Virology* **225,** 156–162, doi:10.1006/viro.1996.0583.

Tanno, F., Nakatsu, A., Toriyama, S., and Kojima, M. (2000). Complete nucleotide sequence of Northern cereal mosaic virus and its genome organization. *Arch. Virol.* **145,** 1373–1384.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

Tordo, N., de Haan, P., Goldbach, R., and Poch, O. (1992). Evolution of negative-stranded RNA genomes. *Semin. Virol.* **3,** 341–357.

Vetten, H. J., Lesemann, D.-E., and Dalchow, J. (1987). Electron microscopical and serological detection of virus-like particles associated with lettuce big vein disease. *J. Phytopathol.* **120,** 53–59.

Wang, Y., McWilliam, S. M., Cowley, J. A., and Walker, P. J. (1994). Complex genome organization in the G_{NS}-L intergenic region of Adelaide river rhabdovirus. *Virology* **203,** 63–72, doi:10.1006/viro.1994.1455.

Wunner, W. H., Calisher, C. H., Dietzgen, R. G., Jackson, A. O., Kitajima, E. W., Lafon, M., Leong, J. C., Nichol, S., Peters, D., Smith, J. S., and Walker, P. J. (1995). Family Rhabdoviridae. *In* "Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 275–288. Springer-Verlag, Vienna.