Virus Genes 31:1, 31–41, 2005

Genome Sequences and Structures of Two Biologically Distinct Strains of *Grapevine leafroll-associated virus* 2 and Sequence Analysis

BAOZHONG MENG,[1,2,*] CAIHONG LI,[2] DARIUSZ E. GOSZCZYNSKI[3] &DENNIS GONSALVES[2,4]

[1]*Department of Molecular and Cellular Biology, College of Biological Science, University of Guelph, N1G 2W1, Ontario, Canada*
[2]*Department of Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, New York, 14456, USA*
[3]*Agricultural Research Council, Plant Protection Research Institute, Private Bag X 134 0001, Pretoria, South Africa*
[4]*USDA-Pacific Basin Agricultural Research Center, 99 Apuni St., Suite 204, Hawaii, 96720, USA*

Received November 7, 2004; Accepted December 28, 2004

Abstract. *Grapevine leafroll-associated virus 2* (GLRaV-2), a member of the genus *Closterovirus* within *Closteroviridae*, is implicated in several important diseases of grapevines including ''leafroll'', ''graft-incompatibility'', and ''quick decline'' worldwide. Several GLRaV-2 isolates have been detected from different grapevine genotypes. However, the genomes of these isolates were not sequenced or only partially sequenced. Consequently, the relationship of these viral isolates at the molecular level has not been determined. Here, we group the various GLRaV-2 isolates into four strains based on their coat protein gene sequences. We show that isolates ''PN'' (originated from *Vitis vinifera* cv. ''Pinot noir''), ''Sem'' (from *V. vinifera* cv. ''Semillon'') and ''94/970'' (from *V. vinifera* cv. ''Muscat of Alexandria'') belong to the same strain, ''93/955'' (from hybrid ''LN-33'') and ''H4'' (from *V. rupestris* ''St. George'') each represents a distinct strain, while *Grapevine rootstock stem lesion-associated virus* (GRSLaV), a virus originally thought to be a new virus that is distinct from all known closteroviruses, should be considered a fourth strain of GLRaV-2. Through cloning the 5′ terminal region of isolate ''94/970'', we determined the complete genome sequence for strain ''PN''. Using an RT-PCR-based strategy, the entire genome of the severe strain ''93/955'' was also sequenced. The genomes of strains ''PN'' and ''93/955'' differ by 10.5% and the differences are unevenly distributed. Sequence analyses using multiple genomic regions confirm the proposition that GRSLaV is a strain of GLRaV-2 rather than a distinct virus. The possibility that GLRaV-2, either acting alone or together with a different virus, may actually cause other diseases totally different from the typical leafroll is discussed.

Key words: *Closteroviridae*, coat protein, genome sequencing, GLRaV-2, grapevine

Introduction

As one of the most important viral disease complexes of grapevines, Grapevine leafroll (GLR) occurs worldwide and causes significant damage to the grape and wine industries around the world [1]. Typical symptoms of GLR include downward curling and discoloration of fully expanded leaves of infected vines toward the end of the growing season. Infection with GLR reduces the yield, and lowers the quality, of grape berries [2]. The etiology of GLR seems to be very complex and has not been definitively resolved. To date, nine serologically distinct viruses, designated *Grapevine leafroll-associated virus* (GLRaV)-1 to 9, have been associated with the disease complex [3,4]; additional viruses are likely to be associated with the

*Author for all correspondence:
E-mail: bmeng@uoguelph.ca

complex in the future. These GLRaVs belong to the family *Closteroviridae* [4,5]: GLRaV-2 is a member of *Closterovirus*, a genus typified by *Beet yellow virus* (BYV, 6) and containing *Citrus tristeza virus* (CTV, 7) and other aphid-transmitted members; GLRaV-1, GLRaV-3 (type member), GLRaV-4, GLRaV-5, GLRaV-6, and GLRaV-8 belong to the newly erected genus *Ampelovirus*; while GLRaV-7 remains as an unassigned member [5].

As the second virus being identified in grapevines with GLR symptoms, GLRaV-2 has long been implicated in the GLR disease complex [5,8,9]. More recently, GLRaV-2 was also reported to be associated with graft-incompatibility syndromes among certain rootstock and scion varieties of grapevines in several countries [10–12]. Furthermore, a virus with a genome structure resembling that of GLRaV-2 was reported to be associated with the quick decline of young vineyards in California where the popular table grapevine variety ''Redglobe'' died 1–2 years after being grafted onto such rootstocks as 5BB, 5C, 3309C, and 1103P [13, 14].

Numerous isolates of GLRaV-2 have been detected in different genotypes of grapevines from different geographic regions. Goszczynski et al. [15] obtained two isolates of GLRaV-2 via sap transmission of the viruses from the grapevine vareities ''Muscat of Alexandria'' and ''LN-33'' to the herbaceous host *Nicotiana benthamiana*, and named them ''94/970'' and ''93/955'', respectively. These two isolates exhibited different pathological properties. First, although both isolates produce chlorotic local lesions on inoculated leaves of *N. benthamiana*, those caused by ''93/955'' turn into solid necrotic lesions. Second, although both isolates elicit vein clearing on non-inoculated leaves emerged after virus inoculation, which changed to vein necrosis, the speed of the transition from chlorosis to necrosis is much faster in the case of ''93/955'' than in the case of ''94/970'' [15]. Third, under the greenhouse conditions in New York, we observed that *N. benthamiana* plants infected with ''93/955'' collapsed more quickly than those infected with ''94/970''. Thus, ''93/955'' seems to be more virulent than ''94/970'', at least when *N. benthamiana* is concerned [15].

A third isolate was obtained in New York from the grapevine cultivar ''Pinot noir'' that was co-infected with GLRaV-2 and GLRaV-3 [8,16,17] and this isolate is designated here ''PN''. Two more isolates were also obtained from the USA, one from the grapevine cultivar ''Semillon'' (designated ''Sem'') and the other from *Vitis rupestris* ''St. George'' (designated ''H4'') [9,18]. Interestingly, the ''Semillon'' plant was derived from the mother plant used as the standard for the disease Grapevine corky bark from Foundation Plant Services at the University of California at Davis.

The genome of an apparent new closterovirus was recently sequenced from quickly declining ''Redglobe'' grapevines [14]. Based on its relatively low sequence identity to GLRaV-2 and its inability to induce leafroll symptoms on the indicator ''Cabernet franc'', the authors suggested that it was distinct from GLRaV-2 and proposed the name *Grapevine rootstock stem lesion-associated virus* (GRSLaV) for it [13,14]. However, this virus seems to be serologically related to isolate ''H4'' of GLRaV-2 (Martelli 2003). Thus, whether GR SLaV should rather be considered a strain of GLRaV-2 has yet to be defined [4].

Despite the fact that biologically different variants of GLRaV-2 were detected in grapevines, their relationship at the molecular level is by and large unknown. This situation is due to the fact that only partial genome sequences have been determined for only some of the isolates. To date, the 3′-terminal 15-kb and 8.6-kb of the genomes of ''PN'' [8] and ''Sem'' [9] have been sequenced, respectively. Although these two isolates originated from different grapevines (''Pinot noir'' versus ''Semilon''), they are virtually identical, with 99.4% nucleotide (nt) sequence identity when the sequenced genome portions are compared. The coat protein (CP) gene of isolate ''H4'' was also sequenced, which had 88% identity in nt sequence to those of isolates ''PN'' and ''Sem'' [18]. On the other hand, the genomes of isolates ''94/970'' and ''93/955'' have not been characterized at all.

The objectives of this study were: (1) to examine the relationship, at the genome sequence level, among ''93/955'', ''94/970'', and with other GLRaV-2 isolates for which partial genome sequences are available; and (2) to determine the entire genome sequences of the two strains represented by ''PN'' and ''93/955''. We report here on the sequencing of the CP genes of ''93/955'' and ''94/970'', and concluded that these two isolates

differed from each other and that "94/970" was virtually identical to isolates "PN" and "Sem". We also determined the complete genome sequences and structures of these two biologically distinct isolates of GLRaV-2. Strains "93/955" and "94/970" differ by 10.5% in nt sequence and the differences were unevenly distributed across the viral genomes. Based on sequence analyses of all genetic regions, we provide further evidence that GRSLaV should be considered a strain of GLRaV-2, rather than a distinct virus species.

Materials and Methods

Virus Source

Virus sources were *N. benthamiana* plants that were infected with strain "93/955" or "94/970" via rub-inoculation. Greenhouse-grown *N. benthamiana* plants at four to six leaf stage were inoculated with viral extracts made from frozen viral stock using the extraction buffer (0.01 M K_2HPO_4, 0.01 M L-cysteine hydrochloride, 3% nicotine and celite).

dsRNA Isolation, Reverse Transcription, and Polymerase Chain Reaction

DsRNAs were isolated from leaves of virus-infected *N. benthamiana* plants and purified via chromatography on cellulose CF-11 column according to the method of Hu et al. [19]. Isolated dsRNAs were converted into cDNAs with *Moloney murine leukemia virus* (MMLV) reverse transcriptase "Superscript II" (Invitrogen). Reverse transcription was carried out at 42°C for one to three hours with appropriate primers. Resulting cDNAs were amplified through PCR using appropriate pairs of primers (Table 1). The PCR parameters used were as follows: an initial denaturation at 98°C for 2 min, followed by 35–40 cycles of denaturation at 94°C for 10 s, annealing at appropriate temperatures for 30 s, and extension at 70°C for 2 min, and a final 10 min extension at 70°C.

Determination of 5′ Terminal Genome Sequence

DsRNAs derived from "94/970" and "93/955"-infected *N. benthamiana* plants were first polyadenylated with poly(A) polymerase (Invitrogen). Polyadenylated dsRNAs were reverse transcribed using MMLV RT at 42°C. cDNAs derived from isolate "94/970" were amplified with PCR using primers $dT_{(17)}$ and BM99-2 (Table 1), while cDNAs derived from isolate "93/955" were amplified with $dT_{(17)}$ and GLR2-41. To ensure that the 5′ terminal sequence of the viral genome was determined, a different approach, 5′ RACE was performed. Briefly, first-strand cDNA fragments corresponding to the 5′ terminal genomic region were obtained via reverse transcription using primer GLR2-41. Reverse transcription was performed using MMLV RT for 2.0 h at 42°C. The 3′ ends of the cDNA frragments were C-tailed using terminal deoxynucleotide transferase (Promega). The tailing reaction was incubated for 30 min at 37°C, followed by inactivation at 80°C for 10 min. The C-tailed cDNAs were amplified with PCR using $dG_{(14)}$ primer and GLR2-41.

Cloning, Sequencing, and Sequence Analyses

Products obtained from RT-PCR amplifications were cloned into either pCR2.1 (Invitrogen) or pGEM-T vector (Promega). Potential recombinant plasmids were selected based on either restriction digestion analysis or quick screen using PCR, purified with a Miniprep kit from Qiagen, and sequenced on an ABI 373 automated sequencer. Sequences of resulting cDNA clones were assembled using SeqMan (DNAStar). Open reading frames (ORF) were discerned using MapDraw (DNAStar). The overall genome sequences and individual ORFs were compared using MegAlign (DNAStar).

Results

Sequencing of the CP Genes of "94/970" and "93/955" and Comparisons with Other GLRaV-2 Isolates

To determine whether "94/970" and "93/955" were different in genome sequence to each other and to other isolates of GLRaV-2, a genomic region encompassing the CP gene and 51 nts downstream sequence of both isolates was amplified with RT-PCR using primers CP-96F and CP-96R (Table 1) and cloned. Four clones derived from "94/970" and three clones from "93/955" were sequenced.

Table 1. Sequences and genomic locations of primers used to clone the genomes of *Grapevine leafroll-associated virus 2* strains "94/970" and "93/955"

Primers	Sequences (5′ to 3′)	Positions
H3-6/T3F (1)[#]	GACTTATCTAGTCGTGTCCG	4558–4577
TC-3T7R (4)	CACGTGCTTGTTGGAATCA	5250–5232
TC-2/T7F1 (2)	TGTCGCAAGCAAGTGTTCG	5883–5901
HEL-16T3R1 (5)	CGATTAGAAGGTATGACGGG	6881–6862
RP-10/T3A (7)	GTCCCTGCGTGATGATGTGAT	9094–9115
TC-15R3 (10)	ATCCAGATAGCCAGCAACTC	9594–9575
GLR2-1 (8)	AGAAGCAGGCTTTCAACAGC	965–984
BM99-2 (9)	CCAAGTAACAGCGCCCATCC	1591–1572
CP-96F (11)	ATGGAGTTGATGTCCGACAG	14577–14596
CP-96R (12)	CAGATTCGTGCGTAGCAGTA	15226–15207
GLR2-28 (28)	TCTCGTCCTCATGAGCAG	1542–1559
GLR2-29 (29)	CCTTTCGAAAGTACTTGC	4520–4503
GLR2-30 (30)	GTGGATGAGTTGAAATCGG	5197–5215
GLR2-31 (31)	AGTTCAGCGAAAGCAACCA	5933–5915
GLR2-32 (32)	AGACGGTAAGGG TTTTCAC	6828–6846
GLR2-33 (33)	CCAGCCACTGTTCAATG	9149–9133
GLR2-34* (34)	atcctgggcccGTGAAGGCTTTCGAGGAAG	9527–9545
GLR2-35* (35)	atcctgggcccGTTATCACCAGGTTGCCAAG	14622–14603
GLR2-36 (36)	GTAATGGAGAATTACGAAG	15172–15190
GLR2-37 (37)	CTATGTCCCTGTTTGAGCA	12838–12856
GLR2-38 (38)	AGGCGTCAGTTATCACCAG	14630–14612
GLR2-39 (39)	TGAAGGCTTTCGAGGAAG	9528–9545
GLR2-40 (40)	TCCAGAAGCTCGCACAAC	11227–11210
GLR2-41 (41)	ACAAAGGTCCGCATTGGAAG	427–408
955–13(13)	AAGGCTTCATGCACCAC	1316–1300
955–17(17)	TCCACGTTTGGTTCATGC	340–357
955–15(15)	CTCGCACCTTCTTCGGT	1005–989
$dT_{(17)}$	$TACGATGGCTGCAGT_{(17)}$	
$dG_{(14)}$	$GGGGGGGGGGGGGG$	

[#]Numbers in parentheses depict the primer numbers used in Fig 3.
*Small case letters are non-viral sequences that contain the recognition site for *Apa* I.

Sequence analysis showed that the four clones from "94/970" were 99.8–100% identical to one another, while those from "93/955" were 99.3–99.6% identical to one another, suggesting the genetic homogeneity of both viral isolates. When clones derived from these two isolates were compared to each other and to the corresponding regions in other isolates, the following findings were obtained. First, "94/970" was 99% identical in nt sequence to isolates "PN" and "Sem", suggesting that these three isolates were virtually identical. Since the genome of isolate "PN" was the first to be extensively sequenced [8], we propose to use "PN" to represent these three nearly identical isolates. Second, "93/955" differed from all other isolates, with nt sequence difference of 7.4% as compared to "PN", and thus should represent a second strain. The nt sequence of the CP of "H4" was reported previously to differ by 12% from that of "PN" [18]. Furthermore, its aa sequence differed by 6.1% from that of "94/970" and by 6.5% from that of "93/955" (not shown), suggesting that "H4" should be considered a third strain of GLRaV-2. When the aa sequence of the CP of GRSLaV was compared, it had 9.5% difference from those of "PN" and "93/955" and 11.1% difference from that of "H4", suggesting that GRSLaV was closely related to these strains. Furthermore, the C-terminal halves of the CPs were nearly identical among the four strains. An alignment of the aa sequences of these four strains is shown in Fig. 1.

Determination of the 5′ Terminal Genome Sequence of Strain "PN"

Since "94/970" was shown to be identical to "PN" based on sequences of their CP genes, we sought to

```
93.955 CP MELMSDDNLGNLVITDASSLNGVDKKLLSAEIIKMLVQKGAPNDGIEVVFGLLLYALAARTTSPKV
H4 CP     ......E..N.......T.............V..........S.......................
PN CP     ......S..S.....................VV..........E.......................
GRSLaV CP .........SG.....................V..........SQ...T..................

93.955 CP QRADSDVIFSNRFGESNVVVTEGDLKKVLDGCAPLTRFTNKLRTFGRTFTEAYVDFCIAYKHKLPQ
H4 CP     ......L....N...R.IA.......................I..........I.............
PN CP     ...........S...R.....................................................
GRSLaV CP ......I..Q.TY.DKT............E.....N.....................V.....M..

93.955 CP LNAAAELGIPAEDSYLAADFLGTCPKLSELQQSRKMFASMYALKTEGGVVNTPVSNLRQLGRREVM
H4 CP     .....................................................................
PN CP     .....................................................................
GRSLaV CP .....................A...............................................
```

Fig. 1. Multiple alignment of the amino acid (aa) sequences of four strains of *Grapevine leafroll-associated virus* 2. Strains included in the alignment are "PN" from *Vitis vinifera* "Pinot noir" [8], "93/955" [this study], "H4" from *Vitis rupestris* "St. George" [18], and GRSLaV (*Grapevine rootstock stem lesion-associated virus*) from *V. vinifera* "Redglobe" [14]. The aa sequence of the CP of strain "93/955" is shown in the top lines. Identical aa sequences were denoted as dots. Note that the C-terminal halves of the CPs are almost identical.

obtain the 5′ terminal missing genome sequence of "PN" using "94/970" as the viral source. DsRNAs isolated from "94/970"-infected *N. benthamiana* plants were polyadenylated, reverse transcribed in the presence of the $dT_{(17)}$ primer, and resulting cDNAs amplified via PCR using $dT_{(17)}$ and the virus-specific primer BM99-2 (Fig. 2). A cDNA product of ca. 400 bps, along with a smear background containing larger DNA fragments, was obtained and cloned. Twenty-five recombinant clones containing inserts of ca. 400 bps or larger were sequenced and resulting sequences were compared to, and assembled with, that of the previously sequenced 3′ terminal portion of the "PN" genome. The positions of these clones relative to the previously sequenced genomic portion

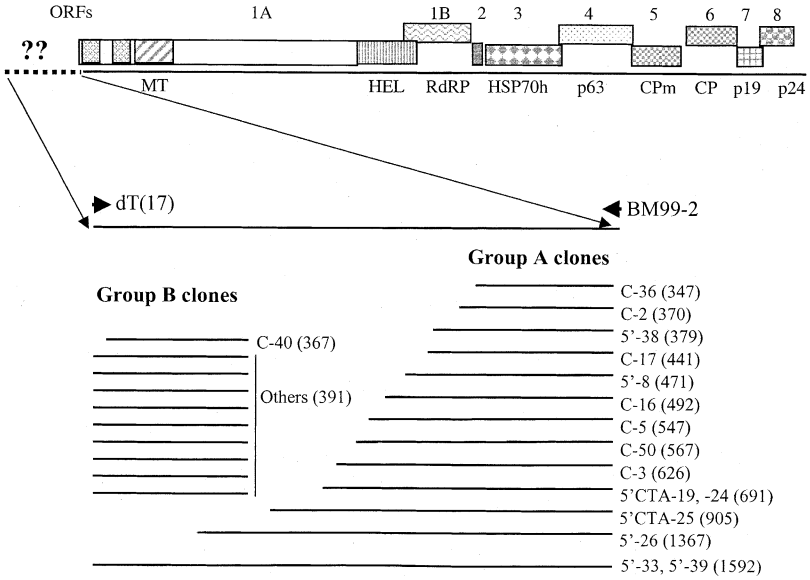

Fig. 2. The strategy used to sequence the 5′ terminal genome region of *Grapevine leafroll-associated virus 2* (GLRaV-2) strain "94/970". The genome structure deduced from the incomplete genome sequence of the isolate "PN" as reported by Zhu et al. [8] was shown on the top. The dotted line indicates the 5′ terminal genome region whose sequence had not been determined prior to this study. The virus-specific primer BM99-2, and a $dT_{(17)}$ primer were used in RT-PCR to amplify the missing genomic region, which resulted in Group A and Group B clones (bottom). Numbers in parentheses indicate the size (in base pairs) of inserts of the cDNA clones. MT: methyltransferase domain; HEL: helicase domain; RdRP: RNA-dependent RNA-polymerase; HSP70h: heat shock protein-70 homologue; CP: the major capsid protein; CPm: the minor capsid protein.

of "PN" are shown in Fig. 2. Overall, these clones belonged to either of two groups. Group A consisted of 15 clones, the inserts of which varied in size, ranging from 347 bps for C-36 to 1,592 bps for 5′-33 and 5′-39 (Fig. 2). As expected, one end of each of these clones mapped to the genomic region based on which the primer BM99-2 was designed. Also as expected, all these clones overlapped the reported "PN" sequence by 98 nts. However, one mismatch was identified in the overlapping region. Unexpectedly, the other ends of these clones extended the reported sequence in the 5′ direction by variable lengths, which ranged from 249 to 1,494 nts (Fig. 2).

In contrast, clones from Group B were more uniform in size and sequence of their inserts. For example, nine of the 10 clones from Group B contained inserts of 391 bps that were identical in sequence. Moreover, these nine clones overlapped partially with 5′-26, and in their entirety with 5′-33 and 5′-39. The latter three clones contained the largest inserts among Group A clones (Fig. 2). The only exception was clone C-40, which had an insert of 367 bps (Fig. 2).

Results from this experiment suggested that the 5′ most terminal nucleotide of the virus genome was likely obtained. To further ensure this, 5′ RACE was conducted using C-tailed first-strand cDNAs. As a result, seven clones were selected and sequenced, five of which matched exactly the 5′ terminal sequence of all but one Group B (C-40) clones and the two Group A clones (5′-33 and 5′-39) that had the largest inserts. The two exceptions were clones 5′ 3-1, which was 17 bps shorter, and clone 5′ 1-2, which was 2 bps longer than the other clones (not shown).

After compiling the sequences obtained from this and a previous work [8], the complete genome sequence of the virus was determined. It comprised 16,493 nts and contained nine ORFs (Fig. 2) as suggested by Zhu et al. [8]. The newly obtained sequences extended the previously reported genome sequence to the 5′ terminus by 1,493 nts, which encompassed the 5′ non-coding region NCR and the first 1,388 nts of ORF1A. The 5′ NCR consisted of 105 nts and had a high A/T content (61 A/Ts), a feature of terminal non-coding sequences of viral genomes. The very first nucleotide of the virus genome was an "A". The now complete ORF1A starts at the first initiation codon at nt positions 106–108, ends at position 8916, and potentially encodes a polypeptide of 2936 aa with an estimated Mr of 328 (Fig. 2).

Determination of the Complete Genome Sequence of Strain "93/955"

Initial sequencing showed that the CP genes of "93/955" and "94/970" differed by 8.5% in nucleotide sequence. This result suggested that these two biologically different isolates of GLRaV-2 also differed in their genomes. As the first step in understanding the molecular basis that underlies the biological differences between the two isolates, we decided to sequence the entire genome of "93/955" using an RT-PCR-based step-wise strategy. The steps involved and primers used in each step to obtain overlapping cDNA clones are shown in Fig. 3, while sequences of the primers and their positions relative to the complete viral genome are given in Table 1. In the first step, four additional small regions at nt positions of 965–1591, 4558–5250, 5883–6881, and 9094–9594 were amplified with RT-PCR using primer pairs derived from the genome sequence of "PN", resulting in cDNA clones E, A-4, B-1, and C-1. In the second and third steps, gaps between these initially obtained clones were bridged through RT-PCR using "93/955"-specific primers derived from neighboring cDNA clones (Fig. 3). In the fourth step, the 5′ terminal 427 nts of the virus genome was cloned and sequenced using essentially the same approach as for "94/970". The 3′ terminal region was obtained from poyladenylated dsRNA using RT-PCR with primers GLR2-36 and $dT_{(17)}$. In the last step, the gap between cDNA clones "D-7" and "E" was bridged through RT-PCR using primers 955-17 and 955-15.

Overlapping cDNA clones derived from 14 genomic regions covering the entire viral genome were obtained, sequenced, and their sequences assembled. Consequently, the complete genome of "93/955" was obtained, which comprised 16,494 bps. Similar to "PN", the genome of "93/955" starts with an "A" at its 5′ terminus, contains a 5′ NCR of 105 nts, followed by nine ORFs (Fig. 3). Except that ORF4 of "93/955" was three nts longer than its counterpart in "PN", all other ORFs had the same number of nucleotides as their

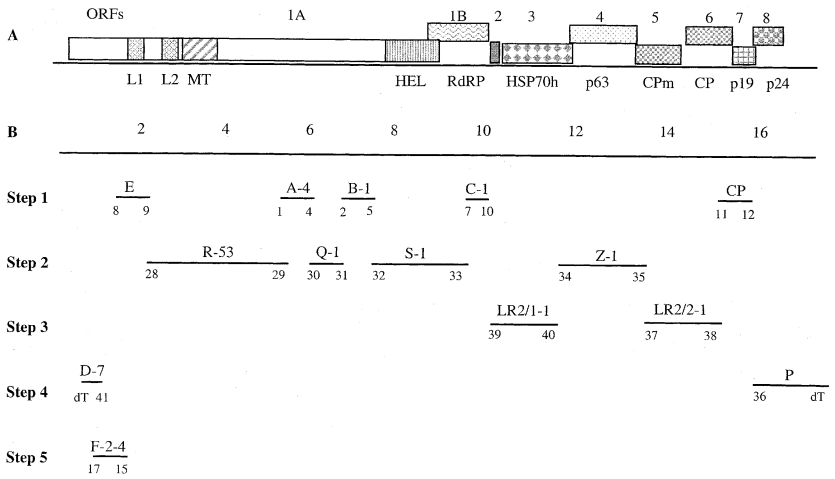

Fig. 3. The genome structure (A) and the strategy (B) used to sequence the genome of *Grapevine leafroll-associated virus 2* strain "93/955". Refer to the legend of Fig. 2 for full names of various gene products. The pair of primers used to obtain each of the overlapping cDNA clones of the virus genome is given as Arabic numerals underneath each clone and are also given in parenthesis in Table 1.

counterparts in "PN". Also like "PN", two putative papain-like protease domains, designated here as "L-1" and L-2", were identified at aa 484–573 and 758–847, respectively, in the polypeptide deduced from ORFlA (Fig. 3).

Comparative Analysis of the Genome Sequences of "93/955", "PN" and GRSLaV

When compared in their entirety, the genomes of "93/955" and "PN" were more closely related to each other (89.5%) than to that of GRSLaV (73.0%, Table 2). The differences in genome sequences were unevenly distributed. At the nt sequence level, the 5′ NCRs of the three viruses were most conserved, with sequence identities ranging from 96.2% between "PN" and "93/955", to 85.7% between "93/955" and GRSLaV, and to 83.8% between "PN" and GRSLaV. In contrast, ORFlAs were the least conserved, with nt sequence identities of 70.6–87.9% among them. At the aa sequence level, ORF1B (RdRP), ORF5 (CPm), and ORF6 (CP) were the most conserved, with identities greater than 89%. It is worth noting that the CPs of the three viruses were closely related, with aa identities of 90–96% among them, suggesting that they were serologically related. In contrast, ORF1A and ORF2 seemed to be the least conserved (Table 2). The distribution of the differences in ORF1As of these viruses was irregular, with a majority of the differences clustered in two regions that were flanked by highly conserved protein domains (Fig. 4). For example, the first variable region encompassed aa 1–852 and was followed by the methyltransferase (MT) domain; the second variable region was located at aa 1701–2400 and was flanked by the MT and the helicase (HEL) domains (Fig. 4). When the two variable regions were excluded from comparison, identities of the remaining aa sequences of the ORF1As rose to 88.4–96.5% among the three viruses (not shown).

Discussion

In this study, we first cloned and sequenced the CP genes of two isolates of GLRaV-2, "93/955" and "94/970". Based on comparison with the CP sequences of previously reported isolates, the relationship of all the GLRaV isolates that have been sequenced was determined. These isolates were grouped into three strains: "PN" (including "PN", "Sem" and "94/970"), "93/955", and "H4". This proposed grouping may be used as a framework for classification of newly described isolates of GLRaV-2 in the future. Based on the fact that "94/970" is virtually identical to "PN", we obtained the 5′-terminal 1,494 nts of the genome sequence of "PN' using "94/970" as the

Table 2. Sequence comparisons of two strains of *Grapevine leafroll-associated* virus 2 and *Grapevine rootstock stem lesion-associated virus* (GRSLaV)

	"93/955" versus "PN"		"93/955" versus GRSLaV		"PN" versus GRSLaV	
	NT	AA	NT	AA	NT	AA
Overall	89.5	–	72.9	–	73.0	–
5′ NCR (105)	96.2	–	85.7	–	83.8	–
ORF1A (106–8916)	87.9	89.2	70.8	72.3	70.6	72.2
ORF1B (RdRP, 8789–10294)	92.4	95.2	78.8	93.0	79.6	94.6
ORF2 (10358–10528)	93.0	91.2	73.7	71.9	72.5	71.9
ORF3 (HSP70h, 10544–12343)	89.7	93.8	74.1	82.7	73.8	82.8
ORF4* (12270–13928)	91.7	95.1	74.1	80.4	73.1	79.8
ORF5 (CPm, 13837–14508)	91.7	96.4	78.1	89.3	78.1	89.3
ORF6 (CP, 14579–15175)	92.1	96.0	76.4	90.0	77.1	90.9
ORF7 (15175–15660)	95.1	96.9	75.9	78.4	76.1	76.5
ORF8 (15662–16279)	89.8	89.3	76.5	77.7	77.0	80.6
3′ NCR (16280 to end)	91.8	–	75.9	–	78.2	–

Listed are percent identities of the entire genomes, their non-coding regions (NCR), and the open reading frames (ORF).
NT, nucleotide sequence; AA, amino acid sequence; –, not applicable.
*ORF4 of "93/955" is three nucleotides longer than that of "PN".

source material for cloning. Compiled with the previously reported 3′ terminal genome sequence, the entire genome of strain "PN" has now been sequenced, which is 16,493 nts in length. Furthermore, using an RT-PCR-based strategy, the entire genome of strain "93/955" has also been sequenced. The genome structures of both strains resemble that of BYV, the type member of the *Closterovirus* genus, and contain nine ORFs. Unlike BYV but similar to CTV, both GLRaV-2 strains possess two leader protease domains in the polypeptides deduced from ORF1A. These two strains differ by 10.5% in nt sequence and the differences are unevenly distributed. The 5′ NCR, ORF1B, ORF2, ORF6 and ORF7 are the most conserved, whereas ORF1A is the least conserved.

An apparently new closterovirus was sequenced from the grapevine "Redglobe" exhibiting quick decline syndrome in newly replanted vineyards, which was tentatively named *Grapevine rootstock stem lesion-associated virus* (GRSLaV, 14). The genome of GRSLaV is 73% identical to those of "PN" and "93/955". We propose that GRSLaV be regarded as a strain of GLRaV-2 rather than a distinct virus species based on the following reasons. First, the genome structure of GRSLaV is identical to those of "PN" and "93/955". Second, the CP and CPm of GRSLaV are closely related to those of "PN" and "93/955", with aa identities of 89.3% and ca. 90% respectively. The high levels of aa sequence identity should warrant serological relatedness among the three isolates. In support of this assertion, virions of GRSLaV were recognized by 18 monoclonal antibodies produced against the CP of GLRaV-2 H4 [4]. Third, the RdRP of GRSLaV is also highly similar to the homologues in "PN" and "93/955", with aa identities of 93–94.6%. This level of sequence identity falls within the boundaries set recently to define members within the same virus species [20]. Lastly, similar or even lower levels of nt sequence identities have been reported for isolates of CTV [21], also a definitive species of the genus *Closterovirus*.

We felt that the 5′ terminus of the GLRaV-2 genome has been reached for the following reasons. First, two independent strategies were used to clone the 5′ terminal region of the virus genome, and results from both experiments point to the same conclusion. Among the two groups of clones obtained using the strategy of tailing the dsRNAs with polyadenylate, Group A clones likely had resulted from specific binding of primer BM99-2 to the dsRNAs, while Group B clones had resulted from nonspecific binding of the same primer to a genome position that was 371-391 nts downstream of its 5′ terminus. Examination of the newly obtained nt sequence at this position revealed the presence of such imperfect annealing site, where 13

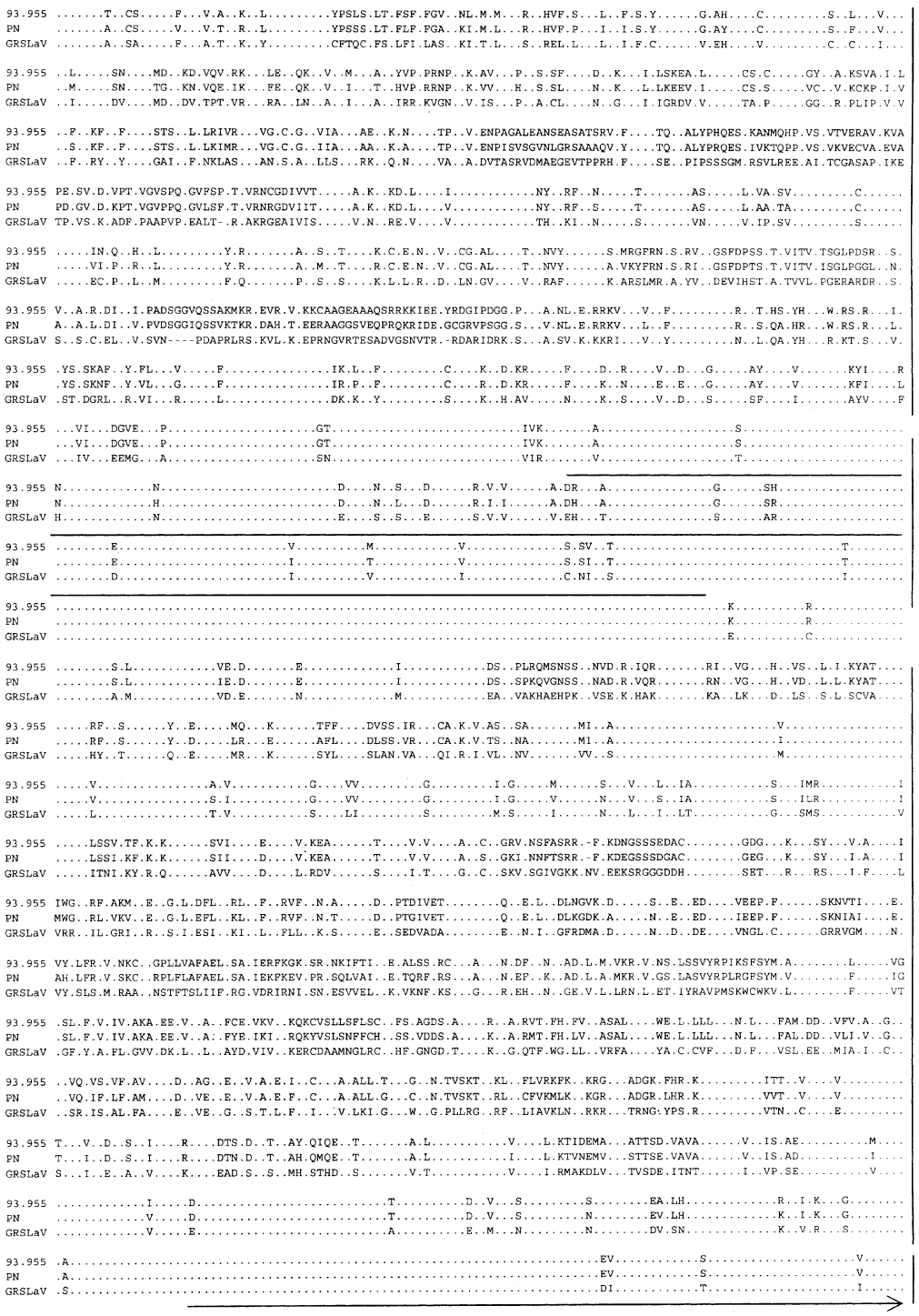

Fig. 4. Alignment of amino acid (aa) sequences of the polypeptides deduced from open reading frame (ORF) 1A of *Grapevine leafroll-associated virus 2* strains "PN" and "93/955" and *Grapevine rootstock stem lesion-associated virus* (GRSLaV). Two variable regions, V1 and V2 at aa 1–852 and 1701–2400, respectively, were identified, which were flanked by the highly conserved methyltransferase (MT) and the helicase (HEL) domains (underlined). Due to space limitation, only the amino-terminal ends of the HEL domains are shown.

of the 20 nts matched the sequence of primer BM99-2 (not shown). The low annealing temperature (42°C) used in RT and PCR permitted such imperfect binding and subsequent amplification. Second, the 5′ NCR of GLRaV-2 is similar in size to those of viruses within the same genus. For example, the 5′ NCRs of BYV [6] and CTV [7], two other viruses within the genus *Closterovirus* whose complete genomes are available, both comprise 107 nts.

The family *Closteroviridae* as recently revised comprises three genera: *Closterovirus*, *Crinivirus*, and the newly established *Ampelovirus* [5,22]. Members of *Closterovirus* are transmitted by aphids except for GLRaV-2 whose transmission by a vector is unknown, have a smaller CP, and the CP is immediately downstream of the CPm. Based on phylogenetic analyses of both replication-related and structural proteins, GLRaV-2 seems to be more closely related to BYV than to CTV (not shown). This may suggest that the former two viruses may have diverged from a common ancestor more recently compared to the divergence of CTV. It remains intriguing as to how the ancestral virus crossed the species barrier to adapt these very different plants as hosts. Because many viruses of the genus *Closterovirus* are transmitted by aphids, it seems logical to suggest that aphids might have been involved in the process.

It is worth noting that both CTV and GLRaV-2 possess two papain-like leader protease domains while BYV and most other viruses of *Closteroviridae* have only one. It is unknown if both protease domains function during the life cycle of the virus. Pena et al [23] recently demonstrated that only the first protease domain of CTV could functionally replace the L-Pro of BYV for genome amplification. Given the even higher sequence homology between the protease domains of GLRaV-2 and that of BYV, it is likely that the leader protease domains of GLRaV-2, at least L1, would also be able to replace the L-Pro of BYV for its cleavage function. However, this hypothesis needs to be tested.

The diseases caused by GLRaV-2 need to be determined. Despite the fact that GLRaV-2 was initially associated with the GLR syndrome and this apparent association is reflected in its name [4,8], the virus may actually cause diseases that are very different from the typical Leafroll. Moreover, the damage caused by GLRaV-2 may turn to be much greater than is commonly believed. To the present, GLRaV-2 has been implicated in several other disease syndromes afflicting grapevines, which include graft-incompatibility [10–12], and quick decline of newly replanted vineyards [13,14]. These diseases seem to be spreading in many grape-growing regions. In many cases, the graft-incompatibility and decline seem to require grafting of a scion variety onto a certain rootstock. If in the end GLRaV-2 proves to be responsible for diseases that are different from Leafroll, the current name of the virus may prove to be inappropriate and consequently the virus would need a new name.

The possibility that another virus may contribute to the aforementioned diseases cannot be excluded. A possible scenario is that the scion variety carries GLRaV-2 and certain rootstocks carry a second virus. Once these two viruses come together in the same vine through grafting, symptoms of graft-incompatibility ensue. The severity of disease symptoms may vary according to the combination of the scion and the rootstock, or according to the virulence of the viral strains that are involved. Thus, GRSLaV may be a virulent strain of GLRaV-2. Survey for the presence of GLRaV-2 and subsequent sequencing of the viral variants may provide insights into the etiology of the aforementioned diseases. It is interesting to note that the only other virus that was detected in the ''Redglobe'' was RSPaV [13]. This makes one wonder if RSPaV was involved in these diseases.

Acknowledgments

This work was supported in part by USDA/ARS cooperative agreement # 58-1908-4-023 with the USDA-Plant Genetic Resources Unit at Geneva, New York, NSERC Discovery Grant (RG261195-03) and the University of Guelph New Faculty Start-up Fund, both awarded to BM. We thank Weizhou Wang and Dan Burke for reviewing the manuscript.

References

1. Martelli G.P., Saldarelli P., and Boscia D., in Monette P.L. (ed.), *Filamentous Viruses of Woody Plants*. Research Signpost, Trivandrum, India, 1997, pp. 1–9.

2. Goheen A.C., in Pearson R.C. and Goheen A.C. (eds), *Compendium of Grape Diseases*. American Phytopathological Society Press, St. Paul, Minnesota, USA, 1988, pp 52.
3. Boscia D., Greif C., Gugerli P., Martelli G.P., Walter B., and Gonsalves D., Vitis *34*, 171–175, 1995.
4. Martelli G.P., Extended Abstracts of 14th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine, Bari, Italy, 2003, pp. 3–10.
5. Martelli G.P., Agranovsky A.A., Bar-Joseph M., Boscia D., Candresse T., Coutts R.H.A., Dolia V.V., Falk B.W., Gonsalves D., Jelkmann W., Karasev A.V., Minafra A., Namba S., Vetten H.J., Wisler G.C., and Yoshikawa N., Arch Virol *147*, 2039–2044, 2002.
6. Agranovsky A.A., Koonin E.V., Boyko V.P., Maiss E., Frotschl R., Lunina N.A., and Atabekov J.G., Virology *198*, 311–324, 1994.
7. Karasev A.V., Boyko V.P., Gowda S., Nikolaeva O.V., Hilf M.E., Koonin E.V., Niblett C.L., Cline K., Gumpf D.J., Lee R.F., Garnsey S.M., Lewandowski D.J., and Dawson W.O., J Virol *208*, 511–520, 1995.
8. Zhu H.Y., Ling K.S., Goszczynski D.E., and Gonsalves D., J Gen Virol *79*, 1289–1298, 1998.
9. Abou-Ghanem N., Sabanadzovic S., Minafra A., Saldarelli P., and Martelli G.P., J Plant Pahol *80*, 37–46, 1998.
10. Greif C., Garau R., Boscia D., Prota V.A., Fiori M., Bass P., Walter B., and Prota U., Phytopathol Medit *34*, 167–173, 1995.
11. Monis J., and Bestwick R.K., Plant Dis *81*, 802–808, 1997.
12. Bonfiglioli R., Edwards F., and Pantaleo A., Aust NZ Grapegrower Winemaker *476*, 50–54, 2003.
13. Uyemoto J.K., Rowhani A., Luvisi D., and Krag C.R., Calif Agric *55*, 28–31, 2001.
14. Rowhani A., Zhang Y.-P., Golino D.A.,and Uyemoto J.K., 2002. Phytopathology *92* S71 .
15. Goszczynski D.E., Kasdorf G.G.F., Pietersen G., and Tonder H.van, Vitis *35*, 133–135, 1996.
16. Ling K.-S., Zhu H.Y., Drong R.F., Slightom J.L., McFerson J.R., and Gonsalves D., J Gen Virol *79*, 1299–1307, 1998.
17. Ling K.-S., Zhu H.Y., and Gonsalves D., J Gen Virol *85*, 2099–2102, 2004.
18. Abou-Ghanem N., Sabanadzovic S., Castellano M.A., Boscia D., and Martelli G.P., Vitis *39*, 119–121, 2000.
19. Hu J.S., Gonsalves D., and Teliz D., J Phytopathol *128*, 1–14, 1990.
20. Adams M.J., Antoniw J.F., Bar-Joseph M., Brunt A.A., Candresse T., Foster G.D., Martelli G.P., Milne R.G., and Fauquet C.M., Arch Virol *149*, 1045–1060, 2004.
21. Mawassi M., Mietkiewska E., Gofman R., Yang G., and Bar-Joseph M., J Gen Virol *77*, 2359–2364, 1996.
22. Karasev A., Ann Rev Phytopathol *38*, 293–324, 2000.
23. Peng C.-W., Peremyslov V.V., Mushegian A.R., Dawson W.O., and Dolja V.V., J Virol *75*, 12153–12160, 2001.