Journal of General Virology (2001), **82**, 2009–2015. Printed in Great Britain

Complete nucleotide sequence and genome organization of Grapevine fleck virus

Sead Sabanadzovic,[1] Nina Abou Ghanem-Sabanadzovic,[1] Pasquale Saldarelli[2] and Giovanni P. Martelli[2]

[1] Istituto Agronomico Mediterraneo, Valenzano (Bari), Italy

[2] Dipartimento di Protezione delle Piante e Microbiologia Applicata, Universitá degli Studi and Centro di Studio del CNR sui Virus e le Virosi delle Colture Mediterranee, Bari, Italy

The complete nucleotide sequence of Grapevine fleck virus (GFkV) genomic RNA was determined. The genome is 7564 nt in size, excluding the 3′-terminal poly(A) tail, is characterized by an extremely high cytosine content (ca. 50%), and contains four putative open reading frames and untranslated regions of 291 and 35 nt at the 5′ and 3′ ends, respectively. ORF 1 potentially encodes a 215·4 kDa polypeptide (p215), which has the conserved motifs of replication-associated proteins of positive-strand RNA viruses. ORF 2 encodes a 24·3 kDa polypeptide (p24) identified as the coat protein. ORFs 3 and 4 are located at the extreme 3′ end of the viral genome and encode proline-rich proteins of 31·4 kDa (p31) and 15·9 kDa (p16), respectively, of unknown function. Phylogenetic analysis of the viral replicase and coat protein genes showed that GFkV is related to members of the *Tymovirus* and *Marafivirus* genera. Two subgenomic RNAs were present in the GFkV preparations as ascertained by molecular hybridization. The genome organization of GFkV resembles to some extent that of tymoviruses and marafiviruses. However, differences in the biological and epidemiological behaviour, cytopathology and molecular properties (i.e. size of genomic RNA and coat protein, and number of ORFs) support the notion that GFkV is a separate virus belonging in a new genus.

Introduction

Grapevine fleck is a widespread graft-transmissible disease of grapevine (Martelli, 1993) caused by Grapevine fleck virus (GFkV), a phloem-limited non-mechanically transmissible virus with isometric particles ca. 30 nm in diameter (Boscia *et al.*, 1991). GFkV virions have a rounded contour and prominent surface structure with coat protein subunits clustered in pentamers and hexamers (Boulila *et al.*, 1990; Boscia *et al.*, 1991). The viral genome is a positive-sense single-stranded RNA ca. 7400 nt in size and the coat protein consists of a single type of subunit with a molecular mass of ca. 28 kDa. During density-gradient centrifugation virus preparations sediment as two components: T, made up of empty protein shells, and B, containing 35% RNA (Boulila *et al.*, 1990). Infected grapevine cells contain cytopathic structures, called multivesiculated bodies, derived from deranged mitochondria that undergo peripheral vesiculation (Castellano & Martelli, 1984). The viral methyltransferase (MTR) and RNA-dependent RNA polymerase (RdRp) are phylogenetically related to those of members of the genus *Tymovirus* and of *Oat blue dwarf virus* (OBDV) genus *Marafivirus* (Sabanadzovic *et al.*, 2000). This paper reports the complete nucleotide sequence of the GFkV genome and discusses the taxonomic position of the virus.

Author for correspondence: Giovanni Martelli.
Fax: +39 080 544 2911. e-mail martelli@agr.uniba.it

Sequence data from this article have been deposited with EMBL under accession no. AJ309022.

Methods

■ **Virus source and purification.** The virus isolate used in this study was the same as in previous investigations (Boulila *et al.*, 1990; Boscia *et al.*, 1995); it was propagated in the grapevine hybrid LN33 and purified from young roots and petioles as described (Boulila *et al.*, 1990). Viral RNA was extracted from purified virus particles by the SDS–phenol method (Diener & Schneider, 1968) and stored in ethanol at −70 °C until use.

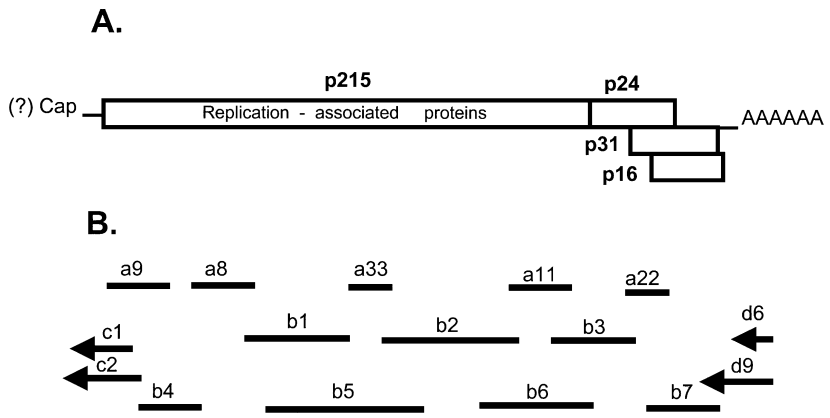

Fig. 1. Schematic representation of the GFkV genome (A) and location of cDNA clones used for its sequencing (B). (A) Boxes represent the ORFs. The putative protein products are indicated above or beside the boxes. (B) Cloning strategy used in GFkV sequencing: a, random primer-generated clones; b, RT–PCR-generated clones using GFkV-specific primers based on sequences of the random-primed cDNA clones; c, 5′-terminal clones generated by RACE–PCR; d, 3′-specific clones generated by dT primers (see Methods).

■ **Cloning and sequencing.** The cloning strategy is shown in Fig. 1. Random primer-generated cDNA was cloned in pGEM-4Z and/or pUC18. The initial clones, encompassing fragments of the virus replication-related proteins and coat protein, were sequenced and used for further cloning, either by primer extension or PCR, to fill the gaps between adjacent clones.

To determine the presence of poly(A) at the 3′ terminus, viral RNA was first submitted to oligo(dT) cellulose chromatography; oligo(dT)-primed cDNAs were then synthesized, cloned and sequenced. The presence of a poly(A) tail was further confirmed by generating dT-primed cDNAs and amplifying the 3′ end with an oligo(dT) and the 3′-proximal GFkV-specific primers Fk19 (5′ GTCCTCCTACACCTCCCTGTCCAT 3′) and Fk3 (5′ CCTCATCCGCGGAGTTATCGAAT 3′), complementary to nt 6609–6632 and 6999–7021 of the GFkV genome, respectively.

The 5′ end nucleotides were determined using a 5′/3′ RACE kit (Roche). Briefly, first strand cDNA was made using AMV RT and the primer FkV5As (5′ AAAGGATGCAGAGCACGAAGCGA 3′; complementary to nt 129–151 of the viral genome). The cDNA preparation was divided into two aliquots to be dA- and dG-tailed separately, according to the manufacturer's instructions. Fragments containing the 5′ end were then generated by PCR using dC or dT primers combined with the internal GFkV-specific primer FkV55As (5′ ACTTGGACAGGG-TGGCGTCAAA 3′), complementary to positions 29–50. Amplified products were cloned using the pGEM-T Easy Vector System (Promega).

All cDNAs, regardless of the vector used, were cloned and subcloned into competent *Escherichia coli* DH5α cells, propagated, and manually sequenced on both strands by the dideoxy chain termination method (Sanger *et al.*, 1977).

■ **Sequence analysis.** Nucleic acid and deduced amino acid products were analysed using the DNA Strider 1.1 program (Marck, 1988). Protein sequences were aligned with CLUSTAL W (Thompson *et al.*, 1994). The GFkV nucleotide and deduced protein sequences were compared with other viral sequences from the GenBank and EMBL databases using the FASTA (Pearson & Lipman, 1988) and BLAST (Altschul *et al.*, 1990) programs. Tentative phylogenetic trees were constructed, and bootstrap analysis performed with the NEIGHBOR, SEQBOOT, PROTDIST and CONSENSE programs of the PHYLIP package (Felsenstein, 1989).

■ **Northern blot.** Viral nucleic acid preparations, extracted separately from T and B fractions, were analysed electrophoretically under semi-denaturing conditions in 1% TBE agarose and stained with ethidium bromide. The gel was incubated in 40 mM NaOH and 2·5 mM EDTA for 20 min and treated with 2 × SSC solution prior to capillary transfer of nucleic acids to Hybond-N+ nylon membranes using 20 × SSC buffer. Membranes were then hybridized at 55 °C with a DIG-labelled cRNA probe complementary to nucleotides 6999–7454, according to the manufacturer's instructions (Roche).

Results and Discussion

The GFkV genome is 7564 nt in size, excluding the poly(A) tail (EMBL accession no. AJ309022). Sequencing was made difficult by the extremely high cytosine content, which led to high compression and termination of readable sequences at the same position in different clones. Over the entire genome length there were two stretches of 13 consecutive cytosine residues (C13), five each of C9, C8 and C7, 12 of C6, and a higher number of C5. A particularly high concentration of C residues occurred in the central part of the genome; for instance, the stretch from nt 2342 to 2441 had 72% C residues. Base ratio was 14·0%A, 16·4%G, 49·8%C and 19·8%T. The cytosine content is higher than that reported for tymoviruses (32–42%) and the marafivirus OBDV (43%) (Koenig, 1988; Edwards *et al.*, 1997) and resulted in an abundance of proline, serine and leucine in virus-coded proteins. As with *Turnip yellow mosaic virus* (TYMV), such an unusual cytosine content could affect the secondary structure of genomic RNA, producing long C-rich single-stranded regions presumably interacting with coat protein (Hellendoorn *et al.*, 1996).

Genome organization

Computer-assisted analysis of the GFkV sequence showed the presence of four main ORFs (Fig. 1). The first two ORFs were in the same frame, separated by a double stop codon.

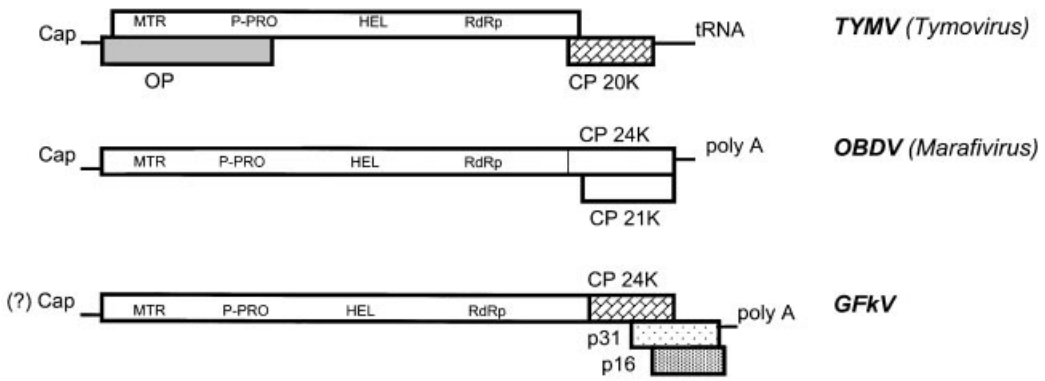

Fig. 2. Comparison of the genome organization of *Turnip yellow mosaic virus* (TYMV; *Tymovirus*), *Oat blue dwarf virus* (OBDV; *Marafivirus*) and *Grapevine fleck virus* (GFkV).

```
GFkV (811aa) ALDCLLVSCSAASGIPKEDLWATLCHIFPPSDL----VSDLGLSTNHLTALAFTYQWLVT
OBDV (736aa) AMDCLLVATEQATRISREALWDCLTATCPDSFLDPKSIAQHGLSTDHFVILAHRFSLCAN
EMV  (773aa) KLNCLLTCFSELSGHSESDLWLSLQSILPDSQLQNPEVSTLGLSTDILTALCFIYHSSVT
TYMV (780aa) TLNCLLSAVSDQTKVSEDHLWESLQTILPDSQLRNEEINSLGLSTEHLTALAHLYNFQAT
              ..*** . .  .  ... **  *    * * *    .   ****. .. *.. .   ..

GFkV  LRSGELVQRHGLLSAPFAFEITHTPPVPPATVGHFALS
OBDV  FHSAEHVIQLGMADATSIFMINHTAGSAG-LPGHFSLR
EMV   LHAPSGVYHYGIASSSTVYVIHYQPGPPP----HFSLS
TYMV  IYSDRGPILFGPSDTIKRIDITHTTGPPS----HFSPG
      . .       *  ..     * . .  .     **.
```

Fig. 3. Alignment of the amino acid sequence around the conserved Cys and His residues (bold) of the (putative) tymo-like proteases of GFkV and some tymo- and marafiviruses. Asterisks and dots indicate identical and similar residues, respectively.

ORF 1 started with an AUG at position 292 and ended with an amber stop codon (UAG) at position 6141. The initiation codon was located in a very favourable translation context (Kozak, 1987), fitting perfectly with the consensus sequence [-GCCG(A)CCAUGG- versus -CCCACCAUGG-]. This ORF encoded a putative polypeptide of 1950 aa with a molecular mass of 215·4 kDa (p215), identified as the replication-associated polyprotein (RP) as it contains the conserved motifs of MTR (Rozanov *et al.*, 1992; Koonin & Dolja, 1993), protease (P-PRO) (Gorbalenya *et al.*, 1991), NTPase/helicase (Gorbalenya & Koonin, 1989) and RdRp of positive-strand RNA viruses (Koonin, 1991; Koonin & Dolja, 1993).

Multiple alignment of GFkV RP with the corresponding polyproteins of tymoviruses and OBDV showed similarities in both sequence and position of conserved motifs (Fig. 2). In particular:

(i) The motifs of the putative MTR domain, a hallmark of the alpha-like plant viruses presumably involved in methylation of the 5′ cap structure (Mi & Stollar, 1991), were found near the N terminus of the GFkV RP (aa 164–341). Although the presence of a cap structure in GFkV RNA was not experimentally demonstrated, the presence of the MTR motifs and of a guanidine in the first position of the viral genomic RNA suggest that GFkV RNA is capped. Comparison of the GFkV MTR with the corresponding domains of tymo- and marafiviruses showed identity ranging from 46% with *Erysimum latent virus* (ErLV; Srifah *et al.*, 1992) to 56% with OBDV (see also Sabanadzovic *et al.*, 2000).

(ii) As with tymoviruses and OBDV (Dreher *et al.*, 2000; Edwards *et al.*, 1997) the MTR domain was followed by an amino acid sequence containing the conserved domain of the putative P-PRO found in a number of alpha-like viruses (Rozanov *et al.*, 1995). TYMV P-PRO, the prototype of this group of proteases was shown to be *cis*-acting and composed of aa 731–885 of the TYMV RP protein (Bransom & Dreher, 1994; Rozanov *et al.*, 1995). All amino acids involved in protease activity and their relative distances (Cys-814 and His-900) were conserved in GFkV RP, which also possessed the CLL and HF/Y motifs conserved in tymoviruses and OBDV (Fig. 3).

(iii) NTPase/helicase motifs I–VI were identified between nt 1043 and 1270 of GFkV RP. The level of identity of this region

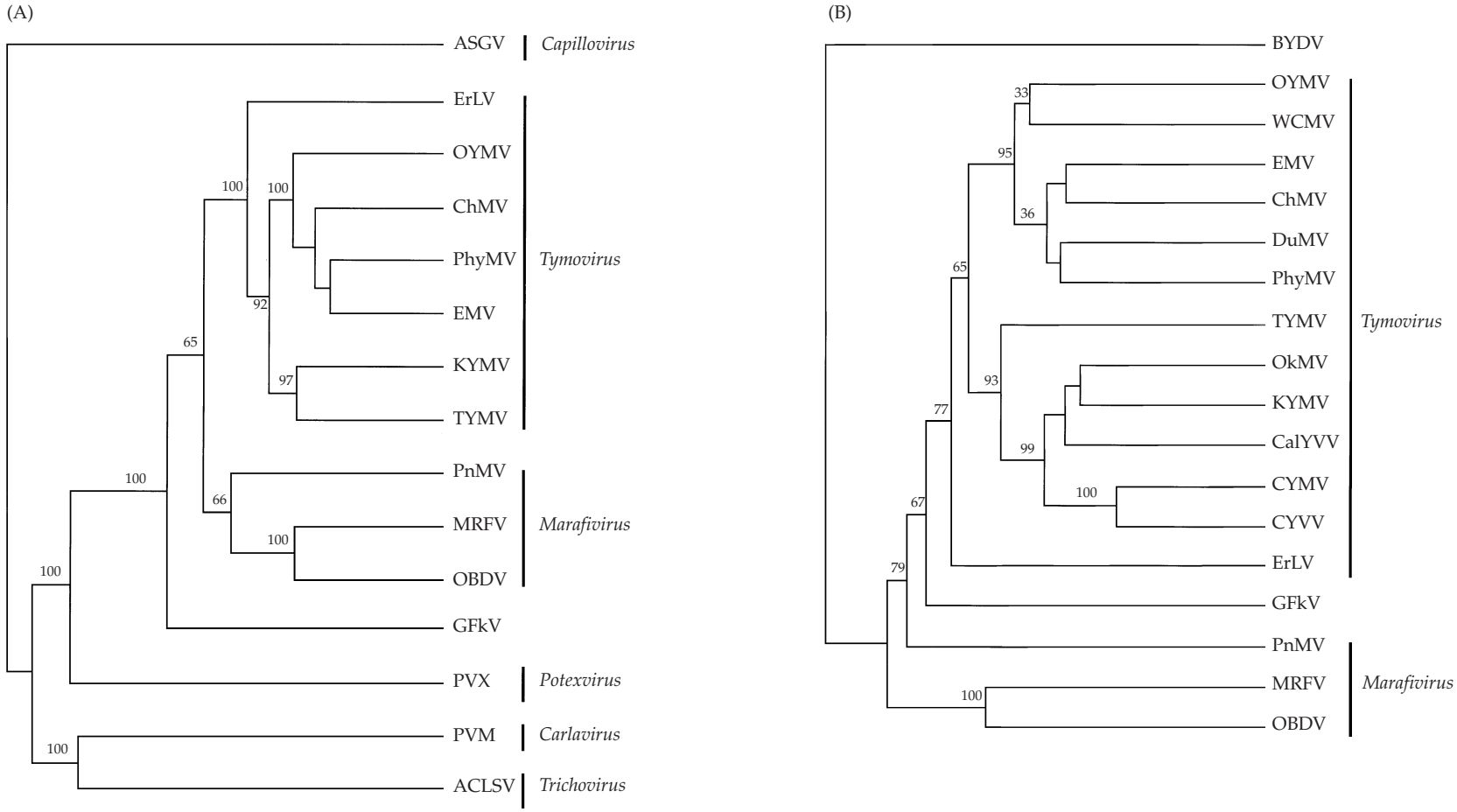

Fig. 4. Phylogenetic analysis of GFkV replicase-associated protein (A) and coat protein (B) with corresponding proteins of some tymo-, marafi-, potex-, carla-, capillo- and trichoviruses. Bootstrap values are shown at branch points. EMBL or GenBank accession numbers of the sequences used in phylogenetic analysis are: *Eggplant mosaic virus* (EMV; J04374); *Kennedya yellow mosaic virus* (KYMV; D00637); *Scrophularia mottle virus* (ScrMV; = *Ononis yellow mosaic virus*; J04375); *Turnip yellow mosaic* (TYMV; X16378); *Physalis mottle virus* (PhyMV; S97776); *Clitoria yellow vein virus* (CYVV; AF035200); *Cacao yellow mosaic virus* (CYMV; X54354); *Chayote mosaic virus* (ChMV; AF195000); *Calopogonium yellow vein virus* (CalYVV; U91413); *Dulcamara mottle virus* (DuMV; AF035634); *Okra mosaic virus* (OkMV; AF035202); *Wild cucumber mosaic virus* (WCMV; AF035633); *Oat blue dwarf virus* (OBDV; U87832); *Poinsettia mosaic virus* (PnMV; AJ271595); *Maize rayado fino virus* (MRFV; AF265566). *Potato virus X* (PVX; X55802), *Potato virus M* (PVM; D14449), *Apple chlorotic leaf spot virus* (ACLSV; AJ243438). *Apple stem grooving virus* (ASGV; D14995) and *Barley yellow dwarf virus*, isolate PAV-129 (BYDV; AF218798) were used as outgroups.

with the NTPases/helicases of tymo- and marafiviruses ranged from 48% with *Physalis mottle virus* (PhyMV; Ranjith-Kumar *et al.*, 1998) to 54% with OBDV.

(iv) The highest level of identity of GFkV RP and the comparable proteins of tymo- and marafiviruses is in the putative RdRp, ranging from 59% with *Ononis yellow mosaic virus* (OYMV; Ding *et al.*, 1989) to 68% with OBDV. This is in line with the previously reported phylogenetic analysis of this region (Sabanadzovic *et al.*, 2000).

A significant difference with tymo- and marafiviruses was represented by the absence from the GFkV genome of the highly conserved 16 nt long subgenomic RNA promoter referred to as the 'tymobox', located near the end of the viral replicase of all sequenced tymoviruses (Ding *et al.*, 1990*a*) and marafiviruses (Edwards *et al.*, 1997; Bradel *et al.*, 2000). Further evidence of the absence of a tymobox was provided by the unsuccessful attempts to amplify this genomic region using tymobox-specific primers. An additional difference with tymoviruses was the absence from the GFkV genome of an ORF with the same size and location of ORF 2 of tymoviruses, which encodes the 'overlapping protein' (OP), a proline-rich putative movement protein (Dreher *et al.*, 2000).

ORF 2 started at position 6367, ended with an opal stop codon at position 7059 and encoded a 230 aa product with a molecular mass of 24·5 kDa (p24), identified as the viral capsid protein (CP). The overall degree of amino acid identity between GFkV CP and CPs of tymoviruses ranged from 23% with OYMV (Ding *et al.*, 1989) to 31% with *Calopogonium yellow vein virus* (Gibbs *et al.*, 1997) and *Cacao yellow mosaic virus* (Ding *et al.*, 1990*b*). Identity with the coat proteins of *Poinsettia mosaic virus* (PnMV) (Bradel *et al.*, 2000) and OBDV was 28 and 29%, respectively. The GFkV CP gene contained the PFQ amino acid triplet conserved in all sequenced tymoviruses and marafiviruses (Edwards *et al.*, 1997; Hammond *et al.*, 1997).

Whereas phylogenetic analysis of RP domains placed GFkV closed to OBDV than tymoviruses (Fig. 4A), analysis of CP sequences placed GFkV in a somewhat intermediate position between tymoviruses and OBDV (Fig. 4B). This was particularly evident when the C-terminal sequences of viral CPs downstream the PFQ motif were used for analysis (not shown).

The third and fourth ORFs overlapped the CP gene towards the 3′ terminus. ORF 3 started at position 6590, ended with an amber stop codon at position 7519, and encoded a 309 aa polypeptide with a molecular mass of 31·4 kDa (p31), rich in proline (33%) and serine (16%). These two amino acids were also highly represented in the putative expression product of ORF 4, with a molecular mass of 15·9 kDa (p16).

Comparison of the polypeptides encoded by ORFs 3 and 4 with proteins from GenBank and EMBL showed similarity with some proline-rich proteins (i.e. extensin-like and hydroxyproline-rich glycoprotein D) and, marginally, with the putative movement protein of tymoviruses. However, the information currently available does not allow any conclusions regarding the potential involvement of these proteins in the intercellular transport of GFkV.

Noncoding regions

The 5′ noncoding region (5′NCR) of GFkV was 291 nt long, much longer than the comparable regions of tymoviruses of which the longest are the 5′NCR of OYMV (171 nt) and OBDV (114 nt) (Ding *et al.*, 1989; Edwards *et al.*, 1997). The GFkV 5′-terminal sequence was determined from five independent RACE–PCR-generated clones and was the same in both dG- and dA-tailed clones. All dA-tailed clones had the 5′ sequence 5′ (A)nGCACAT 3′, whereas all dG-tailed clones had the sequence 5′ (G)nCACAT 3′. It was concluded that the GFkV 5′ end is likely to have the sequence 5′ GCACATTAG 3′.

Analysis of the GFkV 5′NCR showed a prevalence of pyrimidines (235 nt) over purines (56 nt), suggesting that the presence of a pyrimidine-rich region may facilitate translation of the viral RP, being complementary to the sequence present near the 3′ terminus (5′ GGAAG 3′) of wheat 18S ribosomal RNA, as hypothesized for OYMV (Ding *et al.*, 1989) and PhyMV (Ranjith-Kumar *et al.*, 1998). In GFkV, a sequence complementary to 18S ribosomal RNA was found at position 263–267, near the start codon of ORF 1.

The 3′NCR of GFKV RNA was shorter than that of tymoviruses (Ranjith-Kumar *et al.*, 1998) and OBDV (Edwards *et al.*, 1997) and consisted of 35 nt, excluding the poly(A) tail. Sequencing of four independent clones corresponding to the viral 3′ end revealed the presence of a 3′-terminal poly(A) tract from 13 to 24 residues in length. This feature places GFkV closer to OBDV, known to have a polyadenylated 3′ terminus (Edwards *et al.*, 1997), than to tymoviruses, whose genome terminates with a tRNA-like structure (Dreher *et al.*, 2000).

Subgenomic RNAs

Northern blot analysis showed that the top component of GFkV contains at least two subgenomic RNAs with an estimated size of ca. 1300 and 1000 nt, respectively, whereas the bottom fraction contained both genomic and subgenomic RNAs (Fig. 5). The presence of subgenomic RNA in purified virus preparations has been reported for tymoviruses and marafiviruses (Bradel *et al.*, 2000).

Thus, GFkV resembles members of the *Tymovirus* and *Marafivirus* genera in particle structure and outward appearance, physico-chemical and hydrodynamic properties, and partially in intracellular behaviour (Castellano & Martelli, 1984; Boulila *et al.*, 1990; Dreher *et al.*, 2000; Edwards, 2000). In addition, its RP and CP are phylogenetically related, though to different extents, to representatives of both genera (see also Sabanadzovic *et al.*, 2000). By contrast, there are a number of biological and molecular traits that separate GFkV from members of the above two genera. Differences with tymo-

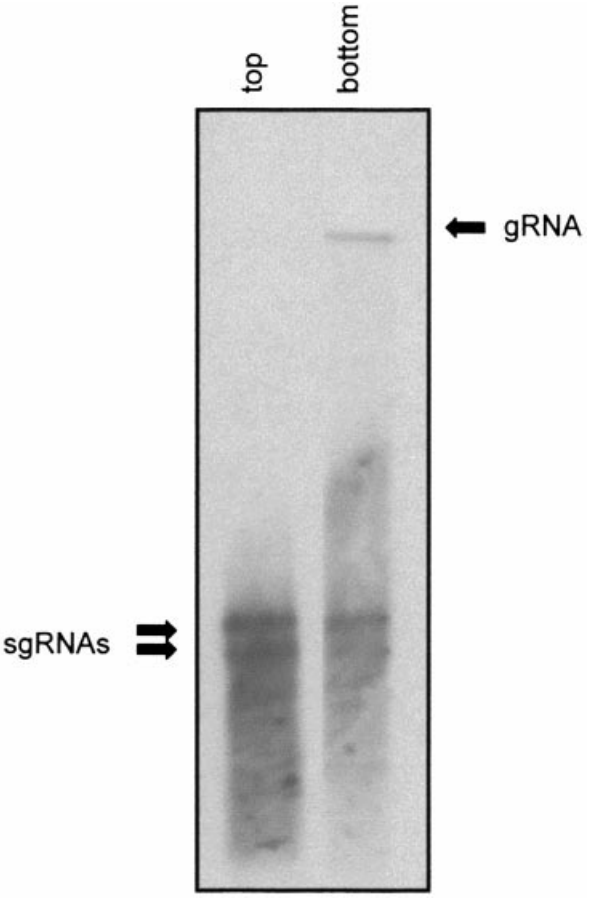

Fig. 5. Northern blot hybridization of nucleic acids extracted from purified viral preparations. The presence of at least two subgenomic RNAs (sgRNAs) is evident in the top and bottom fractions. The bottom fraction contains also GFkV genomic RNA (gRNA).

viruses lie in restriction of GFkV to the phloem, lack of mechanical transmissibility and of a known vector, cytopathology (vesiculation affects mitochondria rather than chloroplasts), genome size (7·5 kb versus 6·3 kb), CP size (28 kDa versus 20 kDa), genome organization, localization and number (4 versus 3) of ORFs, and type of 3′-terminal structure [poly(A) tail versus tRNA-like structure].

Differences with marafiviruses reside in the lack of a known vector, natural host range (*Vitis* versus *Gramineae*), cytopathology (marafiviruses apparently do not induce organellar vesiculation), genome size (7·5 kb versus 6·5 kb), number and size of CP subunits (one with a molecular mass of 28 kDa versus two with molecular masses of 24 and 21 kDa, respectively), genome organization and number of ORFs (4 versus 2). Thus, the general biology and molecular properties of GFkV do not fit those characterizing either the genus *Tymovirus* or the genus *Marafivirus*. The differences are significant enough to support the notion that GFkV should be the representative of a diverse new genus in the *Tymovirus* lineage (Koonin & Dolja, 1993).

The authors wish to thank Dr R. W. Hammond, ARS-USDA, Beltsville, MD, USA, for providing the sequence of *Maize rayado fino virus* prior to publication.

References

Altschul, S. F., Stephen, F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410.

Boscia, D., Martelli, G. P., Savino, V. & Castellano, M. A. (1991). Identification of the agent of grapevine fleck disease. *Vitis* **30**, 97–105.

Boscia, D., Elicio, V., Savino, V. & Martelli, G. P. (1995). Production of monoclonal antibodies to grapevine fleck virus. *Plant Pathology* **44**, 160–163.

Boulila, M., Boscia, D., Di Terlizzi, B., Castellano, M. A., Minafra, A., Savino, V. & Martelli, G. P. (1990). Some properties of a phloem-limited non mechanically-transmissible grapevine virus. *Journal of Phytopathology* **129**, 151–158.

Bradel, G. B., Preil, W. & Jeske, H. (2000). Sequence analysis and genome organization of Poinsettia mosaic virus (PnMV) reveal closer relationship to marafiviruses than to tymoviruses. *Virology* **271**, 289–297.

Bransom, K. L. & Dreher, T. W. (1994). Identification of the essential cysteine and histidine residues of the turnip yellow mosaic virus protease. *Virology* **198**, 148–154.

Castellano, M. A. & Martelli, G. P. (1984). Ultrastructure and nature of vesiculated bodies associated with isometric virus-like particles in diseases grapevines. *Journal of Ultrastructural Research* **89**, 56–64.

Diener, T. O. & Schneider, I. R. (1968). Virus degradation and nucleic acid release in single phenol systems. *Archives of Biochemistry and Biophysics* **124**, 401–412.

Ding, S. W., Keese, P. & Gibbs, A. (1989). Nucleotide sequence of the ononis yellow mosaic tymovirus genome. *Virology* **172**, 555–563.

Ding, S. W., Howe, J., Keese, P., MacKenzie, A., Meek, D., Osorio-Keese, M., Skotnicki, M., Srifah, P., Torronen, M. & Gibbs, A. J. (1990*a*). The tymobox, a sequence shared by most tymoviruses: its use in molecular studies of tymoviruses. *Nucleic Acids Research* **18**, 1181–1187.

Ding, S. W., Mackenzie, A., Torronen, M. & Gibbs, A. (1990*b*). Nucleotide sequence of the virion protein gene of cacao yellow mosaic tymovirus. *Nucleic Acids Research* **18**, 5886.

Dreher, T. W., Haenni, A. L. & Mayo, M. A. (2000). Genus *Tymovirus*. In *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*, pp. 964–969. Edited by 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 & R. B. Wickner. San Diego: Academic Press.

Edwards, M. C. (2000). Genus *Marafivirus*. In *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*, pp. 769–773. Edited by 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 & R. B. Wickner. San Diego: Academic Press.

Edwards, M. C., Zhang, Z. & Weiland, J. J. (1997). Oat blue dwarf marafivirus resembles the tymoviruses in sequence, genome organization, and expression strategy. *Virology* **232**, 217–229.

Felsenstein, J. (1989). PHYLIP–Phylogeny Inference Package, version 3.2. *Cladistics* **5**, 164–166.

Gibbs, A. J., Mackenzie, A. M. & Abdul-Samad, N. (1997). A tymovirus from *Calopogonium mucunoides* in Malaysia is not clitoria yellow vein tymovirus. *Archives of Virology* **142**, 1697–1702.

Gorbalenya, A. E. & Koonin, E. V. (1989). Viral proteins containing the purine-NTP sequence pattern. *Nucleic Acids Research* **17**, 8413–8440.

Gorbalenya, A. E., Koonin, E. V. & Lai, M. M. C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses. *FEBS Letters* **288**, 201–205.

Hammond, R. W., Kogel, R. & Ramirez, P. (1997). Variability of geographically distinct isolates of maize rayado fino virus in Latin America. *Journal of General Virology* **78**, 3153–3159.

Hellendoorn, K., Mat, A. W., Gultyaev, A. P. & Pleij, C. W. A. (1996). Secondary structure model of the coat protein gene of turnip yellow mosaic virus RNA: long, C-rich, single-stranded regions. *Virology* **224**, 43–54.

Koenig, R. (1988). Polyhedral plant viruses with monopartite RNA genomes. Introduction and summary of important properties. In *The Plant Viruses. Polyhedral Virions with Monopartite RNA Genomes*, pp. 1–12. Edited by R. Koenig. New York & London: Plenum Press.

Koonin, E. V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *Journal of General Virology* **72**, 2197–2206.

Koonin, E. V. & Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry and Molecular Biology* **28**, 375–430.

Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **20**, 8125–8132.

Marck, C. (1988). DNA Strider: a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Research* **16**, 1829–1836.

Martelli, G. P. (1993). *Graft-transmissible Diseases of Grapevines. Handbook for Detection and Diagnosis*. Rome: FAO Publication Division.

Mi, S. & Stollar, V. (1991). Expression of Sindbis virus nsP1 and methyltransferase activity in *Escherichia coli*. *Virology* **178**, 429–434.

Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences USA* **85**, 2444–2448.

Ranjith-Kumar, C. T., Gopinath, K., Jacob, A. N. K., Srividhya, V., Elango, P. & Savithri, H. (1998). Genomic sequence of physalis mottle virus and its evolutionary relationship with other tymoviruses. *Archives of Virology* **143**, 1489–1500.

Rozanov, M. N., Koonin, E. V. & Gorbalenya, A. E. (1992). Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. *Journal of General Virology* **73**, 2129–2134.

Rozanov, M. N., Drugeon, G. & Haenni, A.-L. (1995). Papain-like proteinase of turnip yellow mosaic virus: a prototype of a new viral proteinase group. *Archives of Virology* **140**, 273–288.

Sabanadzovic, S., Abou-Ghanem, N., Castellano, M. A., Digiaro, M. & Martelli, G. P. (2000). Grapevine fleck virus-like viruses in *Vitis*. *Archives of Virology* **145**, 553–565.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* **74**, 5463–5467.

Srifah, P., Keese, P., Weiller, G. & Gibbs, A. (1992). Comparisons of the genomic sequences or erysimum latent virus and other tymoviruses: a search for the molecular basis of their host specificities. *Journal of General Virology* **73**, 1437–1447.

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

Received 30 January 2001; Accepted 5 April 2001