Oat Blue Dwarf Marafivirus Resembles the Tymoviruses in Sequences

Michael C. Edwards,*† 2 Zhijun Zhang,† 3 and John J. Weiland*

*United States Department of Agriculture, Agricultural Research Service, Northern Crop Science Laboratory, Fargo, North Dakota 58105-5677; and†Department of Plant Pathology, North Dakota State University, Fargo, North Dakota 58105-5012

Received February 14, 1997; returned to author for revision March 4, 1997; accepted March 21, 1997

The complete nucleotide sequence and genome organization of oat blue dwarf marafivirus (OBDV) were determined. The 6509 nucleotide RNA genome encodes a putative 227-kDa polyprotein (p227) with sequence motifs similar to the methyltransferase, papain-like protease, helicase, and polymerase motifs present in the nonstructural proteins of other positive strand RNA viruses. The 3' end of the open reading frame (ORF) that encodes p227 (ORF 227) also encodes the two capsid proteins: a 24-kDa capsid protein is presumably cleaved from the p227 polyprotein, whereas the 21-kDa capsid protein appears to be translated from a subgenomic RNA (sgRNA). Encoded amino acid and nucleotide sequence comparisons, as well as the OBDV genome expression strategy, show that OBDV closely resembles the tymoviruses. OBDV differs from the tymoviruses in its general biology, in its lack of a putative movement gene that overlaps the replication-associated genes, and in its fusion of the capsid gene sequences to the major ORF. OBDV also possesses a 3' poly(A) tail, as compared to the tRNA-like structures found in most tymoviral genomes. Due to the strong similarities in genome sequence and expression strategy, OBDV, and presumably the other marafiviruses, should be considered a member of the tymovirus lineage of the alpha-like plant viruses.

INTRODUCTION

Oat blue dwarf virus (OBDV) is a plus strand RNA virus with isometric virions approximately 28–30 nm in diameter. The infecting virus is reportedly confined to the phloem tissues in its plant hosts (Banttari and Zeyen, 1971; Zeyen and Banttari, 1972), but also replicates in the aster leafhopper (Macrostele fascifrons), its insect vector (Banttari and Zeyen, 1976). Geographic distribution of OBDV is believed to include Europe as well as North America (Brunt et al., 1996). OBDV has been reported to be serologically related to Bermuda grass etched-line virus (BELV) (Lockhart et al., 1985) and either unrelated (Gámez, 1980) or distantly related (Gingery et al., 1982) to maize rayado fino virus (MRFV), the only other known members of the genus Marafivirus. Studies regarding OBDV transmission (Banttari and Zeyen, 1970; Long and Timian, 1971; Timian and Alm, 1973), ecology (Banttari and Moore, 1962; Westdal, 1968; Timian, 1985), and cytopathology (Banttari and Zeyen, 1972) have been described, but the molecular biology of OBDV and the other marafiviruses is poorly understood.

Although the virions of marafiviruses resemble those of tymoviruses, the two genera are regarded as distinct and unrelated. No genome sequence data has been reported for the marafiviruses and the biological properties of these two groups of viruses differ significantly. MRFV and BELV have narrow host ranges confined to the Gramineae (Brunt et al., 1996), while OBDV has a wider host range which includes both monocotyledonous and dicotyledonous plants (Westdal, 1968; Brunt et al., 1996). All three viruses replicate in their leafhopper vectors and none are readily transmitted by mechanical means (Banttari and Zeyen, 1976; Rivera and Gámez, 1986; Brunt et al., 1996; Louie, 1995). Symptoms induced by MRFV and BELV include stunting and chlorotic streaking or striping along veins, while OBDV induces stunting, stiffening of leaves, enations along leaf veins, and a deepening of color in oat (Brunt et al., 1996). In contrast, the known species of tymoviruses induce bright yellow mosaics and often have narrow host ranges limited to dicotyledonous plants. Most tissues become infected and genome replication is believed to occur within vesicles formed at the periphery of chloroplasts in infected cells. Tymoviruses are readily transmitted mechanically as well as in a non-propagative manner by their natural beetle vectors (Gibbs, 1994). Despite the sharp biological contrasts between the marafi- and tymoviruses, we have identified remarkably strong similarities in the nucleotide sequence, genome organization, and expression strategies of these two viral genera. This is the first reported sequence and detailed analysis of the genome of a marafivirus.
MATERIALS AND METHODS
Virus propagation, purification, and preliminary characterization

The isolate used for these studies was originally obtained from a field of oats near Fargo, ND by R. G. Timian (USDA-ARS, retired) and has been maintained in the greenhouse at Fargo by continuous transfers with the aster leafhopper since 1970. Generally, virus propagation work was conducted between the months of October through April. Healthy leafhopper colonies were maintained on Black Hulless barley (Hordeum vulgare) in a growth chamber at 22–23°C with a 16-hr photoperiod. Viruliferous leafhoppers were obtained by allowing a mixture of adults and nymphs to feed for 7 days on OBDV-infected plants then transferring the leafhoppers to healthy barley for an incubation period of 7 days. These viruliferous leafhoppers were allowed to feed on 4-day-old oat seedlings (Avena sativa “Rodney”) for an inoculation access period of 7 days. Inoculated plants were maintained in the greenhouse at 16°C with a 16-hr photoperiod. Plants with symptoms were harvested 20 days after leafhoppers were removed and frozen at −20°C for subsequent virus purification.

OBDV virions were purified essentially in the way D’Arcy et al. (1983) prepared virions of barley yellow dwarf virus. Virion RNA was extracted by a phenol/sodium dodecyl sulfate (SDS) procedure as described (Petty et al., 1988). Prior to electrophoresis, RNAs were denatured with either formaldehyde or glyoxal (Sambrook et al., 1989). Electrophoresis of proteins was carried out with a 16-hr photoperiod. Virion RNA was extracted by a phenol/so- dium dodecyl sulfate (SDS) procedure as described (Sambrook et al., 1989). Western blotting of proteins was performed as described by Weiland and Edwards (1994), except that a 1:10,000 dilution of rabbit anti-OBDV antisera was used.

Synthesis and cloning of cDNA

Prior to the development of a cloning strategy, oligo(dT) cellulose chromatography was used to assay for the presence of a poly(A) tract in the OBDV genomic RNA. Results suggested that full-length OBDV RNA was indeed polyadenylated (not shown). Therefore the “SuperScript plasmid system for cDNA synthesis and cloning” (Life Technologies) was used as described by the manufacturer. Briefly, RNase H-minus MMLV reverse transcriptase and a NotI primer-adapter (5’GACTAGTTC-TAGATCGGAGCGGCCGCCC[T]3’;Sma I site of the PCR reaction were cloned into the pBluescript II SK(+) (Stratagene).

Manual sequencing was accomplished using the standard dideoxy chain termination method of Sanger et al. (1977). All sequence data resulted from sequencing in both directions and using dGTP analogs (dITP or 7-deazadGTP) to resolve compressions. Subcloning and nested deletion protocols were employed to facilitate sequencing (Sambrook et al., 1989). Primers based on the OBDV sequence obtained, as well as universal forward and reverse primers, were used as needed. Direct sequencing of viral RNA to verify 5’-terminal sequence was done by dideoxy chain termination with AMV reverse transcriptase according to protocols supplied by the manufacturer (USB).

Capsid proteins for sequencing were electroblotted to PVDF membranes according to Moos et al. (1988). Protein sequencing was done by automated Edman degradation in a Poron Model 2090 gas phase sequencer. Initially, sequences were aligned and assembled with MicroGenie (Beckman Instruments). Multiple alignments for sequence comparison with other viruses were done using CLUSTAL V (Higgins and Sharp, 1988) as incorporated into DNAsis version 2.1 for Windows (Hitachi Software). GenBank searches were done using the BLAST programs (Altschul et al., 1990).

Abbreviations and GenBank accession numbers for tymovirus sequences mentioned in this report are: crys-
imum latent virus [ELV (PIR No. J01555, JQ1565)], egg-
plant mosaic virus [EMV (J04374)], kennedya yellow mo-
saic virus [KYM (D00637)], ononis yellow mosaic virus
[OYMV (J04375)], turnip yellow mosaic virus-C [TYMV (X16378)], belladonna mottle virus [BDMV (X54529)], cocoa yellow mosaic [CYMV (X54354)], and physalis mottle virus [PhMV (S97776)]. Other virus abbreviations and GenBank numbers are: apple chlorotic leafspot trichovirus [ACLV (M58152)], potato M carlavirus [PVM (X53062)], apple stem grooving capillovirus [ASGV (D14995)], and potato X potexvirus [PVX (M31541)].

In vitro translation

Rabbit reticulocyte lysate was obtained from Ambion, Inc. and in vitro translation reactions were carried out according to the manufacturer's recommendations using [35S]methionine (DuPont NEN) as a label. Products of in vitro translation were analyzed by SDS – polyacrylamide gel electrophoresis in 7.5, 15, or 5–20% gels according to Laemmli (1970). Following electrophoresis, gels were dried and subjected to autoradiography. Molecular weight estimates of OBDV translation products were calculated using the heavy chain of rabbit skeletal myosin (223 kDa), the tobacco mosaic virus (TMV) 183- and 126-kDa translation products, the brome mosaic virus (BMV) 110- and 94-kDa translation products, bovine serum albumin (68 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), and β-lactoglobulin (18 kDa) as markers.

Subgenomic RNA analysis

Total RNA was extracted from infected oat plants at 2–3 weeks following removal of viruliferous leafhoppers. Fresh tissue was powdered in liquid nitrogen and then ground in 1× STE [50 mM Tris – HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl] containing 1% SDS. After two extractions with phenol/chloroform and one with chloroform, nucleic acids were precipitated from the aqueous phase with one-tenth volume of 3 M sodium acetate and two volumes of ethanol. Denaturation of RNAs with glyoxal and their electrophoresis, blotting, and hybridization were as described (Weiland and Edwards, 1994). A riboprobe complementary to 821 nt at the OBDV 3′-terminus (including 23 A residues) was synthesized using SP6 RNA polymerase and [α-32P]UTP according to Weiland and Dreher (1989).

RESULTS AND DISCUSSION

Initial characterization of the capsid proteins

Disruption of purified OBDV followed by denaturing electrophoresis revealed several possible capsid proteins. The most abundant protein was estimated to be 22.2 kDa, while the other two were 25.5 and 26.6 kDa (Fig. 1A). Preliminary evidence obtained prior to nucleic acid sequencing suggested that these three proteins were related and were of viral origin: All three were readily detectable by Western blotting with anti-OBDV antisera and comparative analysis using limited proteolysis of purified proteins clearly indicated that all three proteins shared common peptide sequences (not shown). The relatedness of the 22.2-, 25.5-, and 26.6-kDa proteins was subsequently confirmed by sequence analysis (see below).

Similarly sized capsid proteins have been reported for the other two marafiviruses. Two capsid proteins of about 22 and 25–29 kDa have been reported for both MRFV and BELV (León and Gámez, 1981; Gingery et al., 1982; Falk and Tsai, 1986; Masumi and Izadpanah, 1996), although only a single capsid protein of about 27 kDa was reported for BELV by Lockhart et al. (1985). These proteins appear to correspond to the OBDV 22.2- and 26.6-kDa proteins; a minor MRFV capsid protein comparable in size to the OBDV 25.5-kDa protein was also reported (Falk and Tsai, 1986). As in our results, peptide mapping indicated that all three MRFV proteins shared common peptide sequences (Falk and Tsai, 1986).

Cloning and sequencing of the genomic RNA

We initially estimated the genome size to be about 6.7 kb after formaldehyde denaturing electrophoresis (see also Pring et al., 1973) and later determined through oligo(dT) chromatography that full-length OBDV RNA harbored a poly(A) sequence of unknown length (not shown).
FIG. 2. Complete nucleotide sequence of the OBDV genomic RNA. The amino acid sequence of the predicted 227 K protein is shown below the nucleotide sequence. A 16-nucleotide sequence similar to the tymobox sequence of tymoviruses is designated and italicized. A filled triangle △
designates the putative polymerase/capsid protein protease cleavage site. Predicted amino acid sequences matching the sequence of the amino terminus of the 24 K capsid protein and a V8 protease-generated fragment of the 21 K capsid protein are italicized and underlined. The putative initiator methionine for the 21 K capsid protein is shaded.
The presence of this poly(A) tract was confirmed by the successful isolation of numerous clones ranging up to 5.6-kb in length after cDNA synthesis that had been primed by oligo(dT). A 5.6-kb clone, designated pOBD9-6, was sequenced in its entirety. Partial sequences of 15 other overlapping clones were aligned with this core sequence as well as each other to provide sequence verification. Sequencing of 13 of these cDNA clones in the region corresponding to the viral 3′ terminus revealed the presence of 3′-terminal poly(A) tracts ranging from 15 to 34 nucleotides in length, with an average length of 23 residues.

The possibility that this poly(A) region might not comprise the actual genome terminus warranted further investigation, as barley stripe mosaic hordeivirus has a poly(A) region followed by a 3′ terminal sequence capable of folding into an amino-acylatable tRNA-like structure (Agranovsky et al., 1981). Evidence consistent with the presence of a poly(A) tail at the 3′ end of the OBDV genome was obtained by T. W. Dreher (personal communication). After 3′ terminal labeling with [5′-32P]pCP, OBDV RNA was digested with RNase T1 (GpN); release of a short unique radiolabeled oligonucleotide would be expected for a 3′ sequence other than poly(A). Instead, only high molecular weight digestion products of a range of lengths were detectable, whereas comparable treatment of TYMV and BMV RNAs resulted in release of a radiolabeled pentamer.

Ligation-anchored PCR allowed the subsequent isolation of clones representing that portion of the genome between the 5′-terminus and the 5.6 kb of sequence already obtained. Sequencing of each of five independent clones of this region completed the OBDV sequence. The 5′-terminal sequence was confirmed by direct sequencing of purified genomic RNA using AMV reverse transcriptase. The OBDV genome comprises 6509 nt, excluding the poly(A) tail (Fig. 2). Base content was determined to be 16.2% A, 19.8% G, 42.8% C, and 21.2% U.

Genome organization

A search for all potential open reading frames (ORFs) revealed the presence of a large ORF between nt 115 and nt 6313, accounting for 95% of the genome (Figs. 2 and 3). The putative polyprotein encoded by this ORF has a calculated molecular mass of 227 kDa (p227). Analysis of the predicted amino acid sequence of p227 revealed the presence of methyltransferase (Rozanov, 1992), protease (Gorbalenya et al., 1991), helicase (Gorbalenya and Koonin, 1989; Habili and Symons, 1989), and polymerase (Kamer and Argos, 1984; Koonin, 1991) motifs known to occur in other positive strand RNA vi- ruses. Although unique ORFs encoding the capsid proteins were not present, translation initiation at AUG1581 and AUG5710 of the p227 ORF would produce related proteins of 25.6 and 21 kDa (Fig. 2).

To determine if either or both of these potential ORFs could encode the 22.2-, 25.5-, or 26.6-kDa proteins observed in purified virion preparations, the amino terminus of each of these proteins was sequenced. Surprisingly, the amino terminal sequences of the 25.5- and 26.6-kDa proteins were identical and did not match the predicted amino terminal sequence of the 25.6-kDa protein putatively encoded from nt 5581 to nt 6313. Instead, a match was obtained between the fifteen terminal amino acid residues and the predicted sequence beginning at Leu1835 of the p227 protein (Fig. 2). These results, and the similarity of the peptide maps of these proteins, suggest that the two species are variants of a single protein that has a calculated MW of 24.4 (p24) and that originates from the p227 polyprotein through enzymatic cleavage between Gly1834 and Leu1835 (Fig. 2). Future experiments will be needed to determine if the appearance of two electrophoretic forms of p24 is due to conformational differences (e.g., TYMV, Matthews, 1974) or posttranslational modification (e.g., Du Plessis and Smith, 1981).

The 22.2-kDa protein is presumably produced by trans-

lation of a subgenomic RNA (sgRNA). Indeed, a putative sgRNA of 900–1000 nt was detected in infected plants when using a probe complementary to the 3′-terminus (Fig. 1B). This putative sgRNA was present in infected plants in much smaller amounts than the genomic RNA and was not detected in purified virions. Failed attempts to sequence the 22.2-kDa capsid protein suggested that its N-terminus was blocked, a phenomenon frequently observed with other viral capsid proteins known to be translated from sgRNAs (Gibbs and Harrison, 1976). Subsequent sequencing of a V8 protease-generated fragment of the 22.2-kDa protein identified a 16 amino acid sequence matching amino acids 1954–1969 of the predicted sequence for p227 (Fig. 2). Thus this capsid protein appears to be encoded by the ORF beginning at AUG5710, the only ORF in this reading frame capable of encoding a protein (p21) with a theoretical MW (21,062) comparable to that expected.

Potential ORFs that could encode proteins as large as 14 kDa were identified in addition to the ORFs encoding the p21 and p227 proteins. Although evidence for the expression of these ORFs was not found, their translation
by noncanonical means (internal initiation, frameshift, etc.) cannot be ruled out.

Similarity to other viral sequences: Resemblance to the tymoviruses

OBDV is the first marafivirus to be sequenced; hence no sequence similarities or relationships with other virus groups have been established. The high cytidine content of the genomic RNA (43%) and the genome size of OBDV were reminiscent of tymoviruses. Subsequent searches of GenBank using the BLAST program (Altschul et al., 1990) revealed significant sequence similarity to tymoviruses (TYMV, KYMV, OYMV, and EMV) at the nucleotide level, with identities of 56–67% occurring over stretches of 500–1400 nucleotides. No other viral RNAs were found to have such high identity percentages with OBDV over such large nucleotide tracts. For example, about 57% identity was found between OBDV and PVX in a 225-nt segment encoding the PVX polymerase domain, whereas 78% identity was found between OBDV and TYMV in the same region. Although these sequence similarities were striking, further analysis revealed still more intriguing similarities between OBDV and the tymoviruses.

(a) The tymobox. A hallmark of the tymoviruses is the presence of a highly conserved 16-nt sequence (the ‘tymobox’) 7–9 nt upstream of the transcription initiation site for the capsid protein sgRNA and very near the end of the replicase polyprotein ORF (Ding et al., 1990). The consensus transcriptional initiation sequence is CAAU, where 5’-AAU is the 5’ terminal sequence of the sgRNA. A sequence of 81% identity (13/16 nt) to that of the tymobox is located in an analogous site in the OBDV RNA sequence (Fig. 2). This level of similarity is greater than that observed in wild cucumber mosaic tymovirus, which has four nucleotide substitutions. Furthermore, a sequence (CAAU) corresponding to the tymovirus transcription initiation box was found 15 nt downstream of the putative tymobox sequence, while a similar sequence (CCAU) was located 7 nt downstream. Transcription initiation at either site would result in production of a ∼950-nt sgRNA, which is consistent in size with the putative sgRNA shown in Fig. 1B. The existence of the putative tymobox and initiation box sequences thus provides further evidence that OBDV produces a subgenomic messenger RNA for expression of p21 from the ORF beginning at nt 5710 and that the strategy for expression of this capsid protein is similar to that of tymoviruses.

(b) The replicase-associated polyprotein. After pairwise alignment for maximum matching (Needleman and Wunsch, 1970), sequence identities of 46–47% were observed between the replicase-associated polyproteins (RP) of OBDV and five tymoviruses, while sequence identities of 49–53% were observed among these tymoviruses. A multiple alignment between these RPs and the putative OBDV RP (minus the 24 K capsid protein sequence) is presented in Fig. 4. Significant similarity in both sequence and genome organization is evident. Generally, the highest identities between regions of the OBDV RP and comparable regions of the tymovirus RPs correlated with regions of high sequence conservation among the tymoviruses. Where tymovirus RP sequences diverged, OBDV RP sequence also diverged. The most extensive identity was found to lie in the putative RNA-dependent RNA polymerase domain.

Comparative analyses of amino acid sequences encoded by positive strand RNA viruses have shown that the tymoviruses are most closely related to the capillo-, carla-, potex-, and trichoviruses (Koonin and Dolja, 1993). Multiple alignment of RP sequences of OBDV, five known tymoviruses, and representative capillo-, carla-, potex-, and trichoviruses revealed significantly greater sequence similarity between OBDV and the tymoviruses than between either of these and the other virus groups. The relative degree of this sequence similarity is evident from the dendrogram in Fig. 5. Similar dendrograms, but with higher matching percentages, were produced after multiple alignments of sequences encompassing the methyltransferase, helicase, and polymerase domains individually.

The replicase-associated polyproteins of a number of alpha-like plant viruses (including the tymo, capillo, carla, and trichoviruses) also possess putative papain-like protease domains (Rozanov, 1995). Sequence comparisons revealed that OBDV encodes a similar putative protease (Fig. 4). The TYMV protease, prototype for this group, has been shown to be a cis-acting protease composed of amino acids 731–885 of the TYMV 206-kDa protein (Bran- som and Dreher, 1994; Bransom et al., 1991; Rozanov et al., 1995). The amino acids Cys783 and His869 are required for TYMV protease activity, while Asp739, Asp809, and Ser810 influence cleavage efficiency (Bransom and Dreher, 1994). These residues and their relative positions are conserved in the OBDV RP with the exception of a three amino acid insertion just upstream of the OBDV homolog to the catalytic His869 residue. The putative active site Cys and His residues of the OBDV RP are followed by aliphatic and aromatic amino acids (CLL and HF), respectively, as is the case in the tymovirus protease domain (CLL and HF/Y). This CLL motif was previously unique to the tymovirus protease. Based upon multiple sequence alignments, the protease domains of OBDV and the tymoviruses have greater sequence similarity between them than either has to the related tymovirus-like protease domains of other viruses.

(c) The capsid proteins. The overall degree of sequence identity between the OBDV 21 K capsid protein (CP) and those of the tymoviruses was much lower than that found with the RPs. Sequence conservation among the tymovirus CPs also is not as great as among the tymovirus RPs. Despite the lack of appreciable general sequence similarity, an alignment of OBDV p21 with the CP sequences of eight tymoviruses revealed that many (18/26) of the amino acid positions highly conserved
FIG. 4. Alignment of the putative OBDV 202 K replicase polyprotein (227 K protein minus the 24 K capsid protein) with the replicase polyproteins of five tymoviruses. Symbols below the TYMV sequence reflect the degree of sequence conservation among the tymoviruses listed, while the symbols below the OBDV sequence reflect the degree of sequence conservation among all six sequences. Sequence identity at a given amino acid position is indicated by (!), (*) indicates identity in at least 4 sequences, and (.) indicates identity in at least 3 sequences. Conserved
methyltransferase (MTR) (after Rozanov et al., 1992; Koonin and Dolja, 1993), helicase (HEL) (after Koonin and Dolja, 1993), and polymerase (REP) (after Koonin, 1991) motifs are identified for reference. The putative papain-like protease domain lies between amino acid positions 731 and 885 of TYMV. The C-783 and H-869 residues are required for TYMV protease activity (Bransom and Dreher, 1994; Rozanov et al., 1995). The putative cleavage site positioned between the helicase and polymerase domains is also indicated.
immediately upstream of the cleavage site (Kadare et al., 1995). Mutagenesis experiments with Sindbis virus demonstrated the importance of Gly or the less optimal Ala in position P2 (Shirako and Strauss, 1990), while sequence comparisons showed that relatively small amino acids (Ala, Gly, Ser, or Cys) were always present in position P1 (Kadare et al., 1995). Alignment of the OBDV p227 sequence with that of five tymoviruses revealed similar sequence conservation at the predicted cleavage site, although the sequence around this site is within tymovirus CPs (Jacob et al., 1992) are also conserved in the OBDV CP sequence (Fig. 6). The CP sequences of other isometric viruses such as BMV and CMV do not possess this level of similarity to either OBDV or the tymoviruses. Also unlike BMV and CMV, the amino termini of the OBDV CPs are not rich in basic residues believed to be involved in RNA–protein interactions (Hopper et al., 1984). Basic amino acids are distributed more uniformly over the length of the OBDV CPs, as is the case for tymoviruses (Jacob et al., 1992). Thus, the OBDV CPs more closely resemble those of tymoviruses than those of other icosahedral viruses.

(d) Expression strategy. Translation of OBDV virion RNA in rabbit reticulocyte lysates resulted in the production of at least four major peptides with estimated Mr of 227, 210, 132, and 100 (Fig. 7). All four were clearly detectable after 45 min of incubation, although the 227-kDa protein decreased in abundance after about 60 min. Proteins migrating at approximately 72 and 25 kDa accumulated at later time points. The lack of distinct ORFs corresponding to these proteins coupled with the decreased abundance of the 227-kDa protein over time suggested the use by OBDV of an expression strategy employing polyprotein processing. Additional support for this model was derived from the existence of the putative papain-like protease domain, the presence of putative nonstructural as well as structural protein sequences within a single ORF (ORF 227), and the sequence similarities between OBDV and the tymoviruses, which are known to employ polyprotein processing (Morch et al., 1989).

Maturation of the TYMV 206 K polyprotein into 140-kDa N-terminal and 66-kDa C-terminal products (Bransom et al., 1991; Kadare et al., 1992) occurs via proteolytic cleavage between Ala1259 and Thr1260 (Kadare et al., 1995; Bransom et al., 1996). Comparison of the amino acid sequences surrounding this cleavage site with the analogous sequences of three other tymoviruses and six alphaviruses revealed significant sequence conservation at the two amino acid positions (designated P1 and P2) immediately upstream of the cleavage site (Kadare et al., 1995). Mutagenesis experiments with Sindbis virus demonstrated the importance of Gly or the less optimal Ala in position P2 (Shirako and Strauss, 1990), while sequence comparisons showed that relatively small amino acids (Ala, Gly, Ser, or Cys) were always present in position P1 (Kadare et al., 1995). Alignment of the OBDV p227 sequence with that of five tymoviruses revealed similar sequence conservation at the predicted cleavage site, although the sequence around this site is...
227OBDV- Tymovirus sequence similarities

and peptide-specific antibodies will be needed to verify this model and to determine the kinetics of peptide accumulation.

Relationship to other taxa

OBDV is clearly a member of the alpha-like plant viruses based on several criteria. Member viruses all have capped RNA genomes, produce sgRNAs, and possess at least three conserved protein domains (methyltransferase, helicase, and polymerase) (Goldbach et al., 1991; Rozanov et al., 1992). Signature motifs for the putative methyltransferase domain, a hallmark of this supergroup, are always encoded near the 5'-terminus and are not found in other viruses (Rozanov et al., 1992). Although direct evidence of a cap structure on the OBDV genomic RNA was not shown, the encoded putative methyltransferase is presumably involved in methylation of the 5' cap structure (Mi et al., 1989; Mi and Stollar, 1991).

The genome organization of OBDV is similar to that of the capilloviruses as well as the tymoviruses. Although sequence similarities are much stronger between OBDV and the tymoviruses, both OBDV and the capilloviruses are polyadenylated and encode large polyproteins containing a capsid protein at the C-terminus (Jelkmann, 1995; Ohira et al., 1995; Yoshikawa et al., 1992). Fusion of the nonstructural and capsid proteins, and the suggestion that capilloviruses encode a papain-like protease (Ohira et al., 1995), raise the possibility that capilloviruses

---

**FIG. 7.** Cell-free translation of OBDV RNA. 35S-labeled proteins were loaded onto a 5–20% gradient gel following translation times of 15–120 min, as indicated. No RNA was added to the reaction loaded onto the first lane (–). Lane (M) contains 14C-labeled markers and marker sizes are indicated along the right border. BMV and TMV RNAs were translated as controls and to provide additional protein markers. Although the myosin marker was listed by the manufacturer (Life Technologies) as 200 kDa on the basis of data obtained by Gazith et al. (1970), subsequent sequence data indicate the actual molecular weight is 223 kDa (GenBank Accession No. U32574). Also note that the predicted OBDV 94 K C-terminal cleavage product and the BMV 94K 2a protein both migrate more slowly than the 97K marker (phosphorylase b).

---

**FIG. 8.** Depiction of the putative papain protease cleavage site between the helicase and polymerase cleavage site (Fig 8B). Amino acid positions equivalent to the crucial P1 and P2 positions are both occupied by Gly. A precedent for multiple site cleavage of a polypeptide by a virus-encoded papain-like protease is provided by the animal alphavirus (Strauss and Strauss, 1990).

Cleavage of the OBDV p227 between Ala1211 and Leu1212 would be expected to yield N- and C-terminal products analogous to those of TYMV.

Peptide sequencing data suggested that the OBDV 24 K CP may also be produced via cleavage from a larger precursor. Examination of the amino acid sequence at this predicted CP cleavage site revealed similarities to the predicted helicase/polymerase cleavage site (Fig 8B). Amino acid positions equivalent to the crucial P1 and P2 positions are both occupied by Gly. A precedent for multiple site cleavage of a polypeptide by a virus-encoded papain-like protease is provided by the animal alphaviruses (Strauss and Strauss, 1990).

Cleavage of the OBDV p227 between Ala1211 and Leu1212 would be expected to yield N- and C-terminal products of 133 and 94 kDa, respectively. The 24 K CP could be cleaved from either p227 or the 94-kDa proteins to yield additional products of either 202 or 70 kDa, respectively. Proteins comparable in size to these predicted fragments were generated by translation of OBDV RNA in rabbit reticulocyte lysates, providing evidence consistent with the proteolytic processing of the 227 K polyprotein in the manner described. Proteins close to 70 and 24 kDa in size were discernible at the later time points, suggesting that p24 can be cleaved from either the 227 K or 94 K proteins. Thus, in vitro translation results support a model for OBDV genome expression wherein p227 is proteolytically processed to yield three products: an N-terminal fragment encompassing the methyltransferase, protease, and helicase domains, a fragment harboring the RNA polymerase, and the 24 K CP. Production of the 21 K CP likely occurs via translation of a sgRNA. Further investigations utilizing mutagenesis and peptide-specific antibodies will be needed to verify this model and to determine the kinetics of peptide accumulation.
also may employ proteolytic cleavage to separate polymerase and capsid sequences.

Due to the strong similarities in genome sequence, organization, and expression strategy, OBDV and perhaps other marafiviruses should be considered members of the tymovirus lineage of alpha-like plant viruses. Unlike tymoviruses, OBDV lacks both a stop codon separating the polymerase and capsid protein genes, and a movement gene analogous to the p69 overlapping protein (OP) gene implicated in the movement of TYMV (Bozarth et al., 1992). Lack of an OP homolog is interesting in light of the hypothesis that the tymovirus OP gene arose de novo via “overprinting” of the preexisting tymovirus RP gene (Keese and Gibbs, 1992). Absence of this gene also may be of fundamental importance; it is possible that the reported phloem limitation of OBDV could occur simply as a result of the absence of an OP equivalent. A third distinction is the apparent termination of the OBDV genomic RNA with a poly(A) sequence. Although aminoclaylation may be required for replication of tymoviruses with tRNA-like structures (Dreher et al., 1996), dulcamara mottle tymovirus has been reported to be polyadenylated (Hellendoorn et al., 1996).

Given the disparate biology that distinguishes the marafiviruses from the tymoviruses, the similarities found in genome sequence, organization, and putative expression strategies were unexpected. The intriguing genome similarities and biological dissimilarities between OBDV and the tymoviruses may provide an avenue for resolving key issues concerning virus evolution, replication, host range, and movement.

ACKNOWLEDGMENTS

The authors thank Theo Dreher for his interest and helpful discussions in addition to his assistance in the characterization of the OBDV 3’ end. The technical assistance of Renee McLean and Kerman Alm is greatly appreciated. Peptide sequencing was performed by Leonard Cook of the Biopolymers Service Center, Department of Biochemistry, North Dakota State University.

REFERENCES


mosaic virus RNA: long, C-rich, single-stranded regions. Virology
224, 43–54.
multiple sequence alignments on a microcomputer. Gene
73, 237–244.
bushy stunt virus coat protein sequence determination and its struc-
tural implications. J. Mol. Biol. 177, 701–713.
sequence of the 3′ terminal region of belladonna mottle virus-lowa
(renamed Physalis mottle virus) RNA and an analysis of the relation-
sequence analysis and serology reveal a new plant capillovirus in
parison of the strategies of expression of five tymovirus RNAs by in
turnip yellow mosaic virus protease in Escherichia coli and determination of
the cleavage site within the 206 kDa protein. J. Gen. Virol. 76, 2853–2857.
dependent polymerases from plant, animal and bacterial viruses.
Nucleic Acids Res. 12, 7269–7282.
Koonin, E. V. (1991). The phylogeny of RNA-dependent RNA polymer-
Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-
strand RNA viruses: Implications of comparative analysis of amino
Laemmli, U. K. (1970). Cleavage of structural proteins during the assem-
León, P., and Gámez, R. (1981). Some physicochemical properties of
Lockhart, B. E. L., Khaless, N., Lennon, A. M., and El Maataouii, M.
(1985). Properties of bermuda grass etched-line virus, a new leafhopper-
transmitted virus related to maize rayado fino and oat blue dwarf
branes and transmission of oat blue dwarf virus by Macrosteles
fasciicorns. Phytopathology 61, 1230–1232.
transmission of maize white line mosaic virus and other viruses of
maize. Phytopathology 85, 139–143.
isolated in cesium chloride density gradients. Virology 60, 54–64.
tion of the Sindbis virus RNA methyltransferase activity with the non-
Mi, S., and Stollar, V. (1992). Expression of Sindbis virus nsP1 and
sequencing of proteins electrophoretically separated and transferred to
lytic origin of the 150-kilodalton protein encoded by turnip yellow
tble to the search for similarities in the amino-acid sequence of two
Complete sequence of an infectious full-length CDNA clone of citrus
udder leaf capillovirus: comparative sequence analysis of capillovirus
for one-step cloning of full-length viral cDNA: application to barley
stripe mosaic virus RNAs. Gene 74, 423–432.
Pring, D. R., Zeyen, R. J., and Battarri, E. E. (1973). Isolation and charac-
terization of oat blue dwarf virus ribonucleic acid. Phytopathology
63, 393–396.
Rivera, C., and Gámez, R. (1986). Multiplication of maize rayado fino
virus in the leafhopper vector Dalbulus maidis. Intervirology 25, 76–
82.
the putative methyltransferase domain: a hallmark of the ‘sindbis-
like’ supergroup of positive-strand RNA viruses. J. Gen. Virol. 73,
2129–2134.
tease of turnip yellow mosaic virus: a prototype of a new viral prote-
A Laboratory Manual,” 2nd ed. Cold Spring Harbor Laboratory Press,
Plainview, New York.
5467.
Shirako, Y., and Strauss, J. H. (1990). Cleavage between nsP1 and nsP2 initi-
es the processing pathway of Sindbis virus nonstructural poly-
protein P123. Virology 177, 54–64.
fasciicorns for increased efficiency in virus transmission. Phytopathol-
ogy 63, 109–112.
Tesser, D. C., Brousseau, R., and Vernet, T. (1986). Ligation of single-
stranded oligodeoxyribonucleotides by T4 RNA ligase. Anal. Bio-
Timian, R. G. (1985). Oat blue dwarf virus in its plant host and insect
Troutt, A. B., McHeyzer-Williams, M. G., Pulendran, B., and Nossal,
G. J. V. (1992). Ligation-anchored PCR: A simple amplification tech-
ique with single-sided specificity.Proc. Natl. Acad. Sci. USA 89,
9023–9025.
CDNA: Effects in vitro and in vivo of point substitutions in the
initiation codons of two extensively overlapping ORFs. Nucleic Acids
Res. 17, 4675–4687.
Weiland, J. J., and Edwards, M. C. (1994). Evidence that the αa gene of
barley stripe mosaic virus encodes determinants of pathogenicity to
J. Bot. 46, 1431–1435.
obide sequence of apple stem grooving capillovirus genome. Virology
191, 98–105.
Zeyen, R. J., and Battarri, E. E. (1972). Histology and ultrastructure of